

METHOD FOR  
INCREASING THE CONTENT  
OF  
SULPHUR COMPOUNDS  
AND IN PARTICULAR  
OF  
CYSTEINE,  
METHIONINE AND GLUTATHIONE  
IN PLANTS AND PLANTS  
OBTAINED

Richard DeRose  
Dominique Job  
Michel Droux

-and-

Anne Lappartient

INTERNATIONAL APPLICATION

-with- Search Report

FOR

PCT/FR99/03179 filed December 17, 1999

-with- Twelve (12) Sheets of Drawings

PH-98/080 ..... (5500\*42)

"Express Mail" mailing label  
number EK219465165

Date of Deposit  
- February 22, 2000 -

I hereby certify that this paper or fee is  
being deposited with the United States Postal  
Service "Express Mail Post Office to  
Addressee" service under 37CFR 1.10 on the  
date indicated above and is addressed to the  
Assistant Commissioner for Patents,  
Washington, D.C. 20231  
- Jean Marshall -

(Typed or printed name of person mailing  
paper or fee)

*Jean Marshall*  
(Signature of person mailing paper or fee)

**Procédé pour augmenter la teneur en composés soufrés et notamment en cystéine, méthionine et glutathion chez les plantes et plantes obtenues**

5           La méthionine est le premier acide aminé essentiel limitant chez les plantes, en particulier les légumineuses qui sont une des bases de l'alimentation animale. La cystéine, autre acide aminé soufré, n'est pas un acide aminé essentiel, mais peut être considérée comme un élément limitant pour la nutrition animale puisque la cystéine dérive, chez les animaux, de la méthionine. Dans le maïs, les acides aminés soufrés sont  
10 aussi des acides aminés limitant après la lysine et le tryptophane. En effet les protéines de réserve majoritaires des graines de ces plantes, sont pauvres en ces acides aminés. La surproduction de méthionine et de cystéine dans les graines des légumineuses (soja, luzerne, pois,...) et du maïs aura donc un impact considérable sur la qualité nutritionnelle de ces graines.

15           Jusqu'à présent, l'augmentation de la qualité nutritionnelle des aliments dérivés des graines de légumineuses a été obtenue par complémentation avec de la méthionine libre synthétisée chimiquement. Par exemple, les contenus moyens en méthionine + cystéine du soja et du pois sont de l'ordre de 20 mg par g de protéines. Ce contenu doit être augmenté à une valeur de l'ordre de 25 mg cystéine + méthionine/g de protéines  
20 pour couvrir les besoins alimentaires de l'homme adulte, et à une valeur de l'ordre de 48 mg de cystéine + méthionine/g de protéines pour couvrir ceux des porcs (De Lumen, B.O., Food Technology (1997) 51, 67-70).

          Les techniques de caractérisation des protéines enrichies en acides aminés soufrés et la préparation de plantes transgéniques permettant l'expression de telles  
25 protéines, de manière à augmenter la teneur en acides aminés soufrés de ces plantes, et donc leur valeur nutritive pour l'alimentation animale, donc à diminuer l'apport en méthionine de synthèse, sont maintenant bien connus et décrites dans la littérature([1] Korit, A.A. & al., Eur. J. Biochem. (1991) 195, 329-334 ; WO 98/20133 ; WO 97/41239 ; WO 95/31554 ; WO 94/20828 ; WO 92/14822).

30           L'enrichissement en protéines à fortes teneurs en acides aminés soufrés par une telle approche reste toutefois limité à la capacité des cellules végétales et des plantes à produire lesdits acides aminés soufrés nécessaire à la sythèse de la protéine. En effet, les plantes sur-exprimant une protéine riche en méthionine et cystéine au niveau de leur graine, comme par exemple les lupins exprimant l'albumine 8S, présentent un taux en  
35 méthionine, cystéine libre et aussi glutathion inférieur aux plants témoins ( [2] Tabe , L. & Droux, M., 4<sup>th</sup> Workshop on Sulphur Metabolim, sous presse).

          De même, des peptides riches en acides aminés soufrés présentant une activité antifongique ou antibactérienne ont été identifiés (WO 97/30082, WO 99/02717, WO 9909184, WO 99/24594, WO 99/53053). L'expression de ces peptides dans les plantes

permet d'augmenter la capacité des dites plantes à résister à certaines agressions fongiques ou bactériennes. Là encore, la production de tels peptides dans les plantes reste limitée à la capacité des cellules végétales et des plantes à produire les aminés soufrés nécessaires à la synthèse de ces peptides. En effet, l'expression de ces peptides dans la cellule végétale se fait au détriment du stock de glutathion, considéré comme un réservoir pour la cystéine.

Or, on a constaté que le paramètre limitant d'une telle approche est bien lié à cette capacité à produire de la méthionine ou de la cystéine. Il est donc important de pouvoir modifier dans les plantes cette capacité à produire de la méthionine et de la cystéine en quantités suffisantes pour permettre la production de protéines hétérologues à haute teneur en acides aminés soufrés, c'est à dire de mettre en œuvre une stratégie moléculaire visant à augmenter les taux de cystéine et de méthionine chez les plantes, et plus particulièrement les plantes de culture d'intérêt agronomique.

Chez les plantes, la biosynthèse de méthionine est effectuée à partir de la cystéine, cette même cystéine étant impliquée dans la synthèse du glutathion.

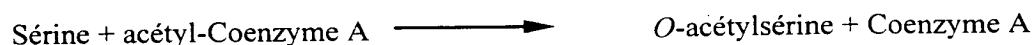
Le glutathion est une forme de stockage du soufre réduit et représente 60 à 70% du soufre organique dans la cellule. Le glutathion joue un rôle important pour les plantes dans la résistance contre les stress oxydatifs et l'élimination des composés toxiques. Il participe ainsi à l'élimination des composés xénobiotiques: des métaux lourds (par exemple) via la formation des phytochélatines et des métallothionines; des herbicides, via l'activité glutathion S-transférase; qui sont toxiques pour la plante, et dans des mécanismes de défenses de la plante contre les micro-organismes. En augmentant la teneur en cystéine d'une plante, et par conséquent sa teneur en glutathion, il est alors possible de moduler la réponse de la plante aux différents stress cités ci-dessus.

Il existe donc à partir de la cystéine, deux voies métaboliques distinctes, l'une pour la préparation de la méthionine, l'autre pour la préparation du glutathion (**Figure 1**) et dont les différentes enzymes impliquées sont rappelées ci-après. Les activités SAT (E1) et OASTL (E2) sont à un carrefour métabolique entre l'assimilation de l'azote et du carbone organique (sérine) et du soufre inorganique (soufre réduit issue de la séquence d'assimilation et de réduction du sulfate, cadre grisé). La cystéine est ensuite incorporée dans les protéines, mais aussi participe à la synthèse du glutathion et de la méthionine. Pour ce dernier acide aminé, la synthèse de son squelette carbonée (*O*-phosphohomoserine) dérive de l'aspartate. L'aspartate est aussi le précurseur à la synthèse de la lysine, de la thréonine et de l'isoleucine. D'autre part est indiquée sur le schéma la présence d'une étape limitante potentielle pour la synthèse de la méthionine, par la régulation au niveau transcriptionnelle de la CGS (cystathionine  $\gamma$ -synthase) ([3] Giovanelli J. in *Sulfur Nutrition and Sulfur Assimilation in Higher Plants*, (1990) pp. 33-48 ; [4] Chiba Y. & al. (1999), *Science*, 286, 1371-1374). La méthionine est le

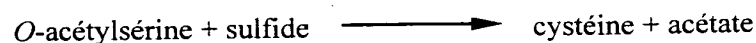
précurseur du SAM (*S*-adenosylméthioninē) impliqué dans la plupart des réactions de méthylations, et du SMM (*S*-méthylméthionine) considéré comme une forme de transport et de stockage de la méthionine ([3]).

Chez les plantes les étapes finales de synthèse de la cystéine impliquent les deux enzymes suivantes :

**E1) Sérine acétyltransférase (EC 2.3.1.30) (SAT):**

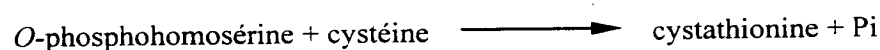


**E2) O-acétylsérine (thiol) lyase (EC 4.2.99.8) (OASTL):**



La synthèse de la méthionine à partir de la cystéine implique une succession des trois enzymes suivantes :

**E3) cystathionine  $\gamma$ -synthase (EC 4.2.99.9) (CGS) :**

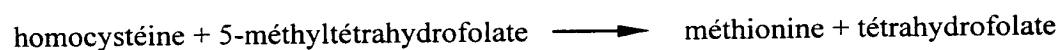


Pi signifie phosphate inorganique.

**E4) cystathionine  $\beta$ -lyase (EC 4.4.1.8) (CBL) :**

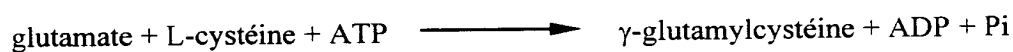


**E5) méthionine synthase (EC 2.1.1.14) (MS) :**

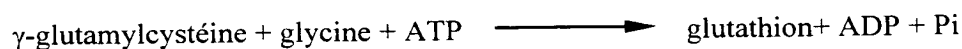


La synthèse du glutathion à partir de la cystéine implique pour sa part une succession des deux enzymes suivantes :

**E6)  $\gamma$ -glutamylcystéine synthétase (EC 6.3.2.2)**



**E7) glutathion synthétase (EC 6. 3.2.3)**



25

Toutes ces enzymes ont été caractérisées et clonées chez les plantes ([5] Lunn, J.E. & al., *Plant Physiol.* (1990) 94, 1345-1352 ; [6] Rolland, N & al., *Plant Physiol.* (1992) 98, 927-935 ; [7] Droux, M. & al, *Arch. Biochem. Biophys.* (1992) 295, 379-390 ; [8] Rolland, N. & al., *Arch. Biochem* (1993) 300, 213-222 ; [9] Ruffet, M.L. & al., *Plant Physiol.* (1994) 104, 597 - 604 ; [10] Ravanel, S. & al., *Arch. Biochem. Biophys.* (1995) 316, 572 - 5584 ; [11] Droux, M.& al, *Arch. Biochem. Biophys.* (1995) 31, 585 - 595 ; [12] Ruffet, M.L. & al., *Eur. J. Biochem.* (1995) 227, 500 - 509 ; [13] Ravanel, S. & al., *Biochem. J.* (1996) 320, 383 - 392 ; [14] Ravanel, S. & al., *Plant Mol. Biol.* (1996) 29, 875 - 882 ; [15] Rolland, N. & al., *Eur. J. Biochem.* (1996) 236, 272 - 282 ; [16] Ravanel, S. & al., *Biochem. J.* (1998) 331, 639-648 ; [17] Droux, M &

al., Eur. J. Biochem. (1998) 255, 235-245; [18 ] May, M.J., Leaver, C.J., Proc. Natl. Acad. Sci. USA (1994) 91, 10059-10063 ; [19 ] Ullmann, P. & al., Eur. J. Biochem. (1996) 236, 662-669 ; [20 ] Eichel, J. & al., Eur. J. Biochem. (1995) 230, 1053-1058).

On sait que pour la synthèse de cystéine, les enzymes E1 et E2 sont présentes dans les trois compartiments de la cellule végétale, c'est-à-dire, les plastes, le cytosol et les mitochondries (5-6, 9, 12). Ces trois enzymes E1 sont dénommées SAT2 et SAT4 pour l'enzyme (putative) chloroplastique, SAT1 pour l'enzyme mitochondriale, SAT3 et SAT3' (SAT52) pour l'enzyme cytoplasmique. Ces attributions des localisations sont basées sur l'analyse des séquences.

Pour les enzymes de la synthèse de la méthionine, la situation est différente puisque les enzymes E3 et E4 sont exclusivement localisées dans les plastes (10-11, 13-14, 16), alors que l'enzyme terminale E5 est dans le cytosol (20).

Les enzymes associées à la voie de biosynthèse du glutathion sont pour leur part localisées à la fois dans le chloroplaste et le cytosol ([21] Hell, R. and Bergmann, L., Planta (1990) 180, 603-612).

L'enzyme E3, de la voie de synthèse de la méthionine, présente un  $K_m$  (concentration en substrat donnant la moitié de la vitesse maximum) de l'ordre de 200  $\mu M$  à 500  $\mu M$  pour la cystéine (10, 16, [22] Kreft, B-D. & al., Plant Physiol. (1994) 104, 1215-1220).

L'enzyme E6, de la voie de synthèse du glutathion, présente aussi un  $K_m$  élevé pour la cystéine, de l'ordre de 200  $\mu M$  [21].

On a maintenant constaté que l'enzyme serine acétyltransférase chloroplastique (**Figure 2**) et à un moindre niveau la SAT mitochondriale étaient inhibées par la cystéine, contrairement à l'enzyme cytoplasmique (**Figure 2**), cette inhibition constituant le facteur limitant essentiel de la synthèse de cystéine dans les cellules végétales, et en aval de la méthionine et du glutathion.

La présente invention consiste donc à augmenter le taux de cystéine et de méthionine synthétisées dans les compartiments cellulaires des cellules végétales, et notamment dans le compartiment chloroplastique. L'augmentation du taux de cystéine, précurseur soufré du glutathion et de la méthionine et ses dérivés, permet avantageusement d'augmenter le taux de méthionine et/ou de glutathion dans les cellules végétales et les plantes, et subséquentement d'améliorer la production de protéines, naturelles ou hétérologues, enrichies en acides aminés soufrés dans les cellules végétales et les plantes, de même que la tolérance des plantes aux différents stress régulés par le glutathion.

Cette augmentation selon l'invention est obtenue en surexprimant une Sérine acétyltransférase (SAT) dans les cellules végétales et les plantes.

La présente invention concerne donc un procédé pour augmenter la production de cystéine, glutathion, méthionine et leurs dérivés soufrés, par les cellules végétales et

les plantes, ledit procédé consistant à surexprimer une SAT dans les cellules végétales, et des plantes contenant les dites cellules végétales.

La SAT surexprimée peut être constituée par toute SAT, qu'elle soit d'origine végétale, notamment SAT2 ou SAT4, SAT1, SAT3, SAT3' (SAT52), ou de toute autre  
 5 origine, notamment bactérienne, sous une forme native ou mutée ou délétée d'un fragment, et fonctionnelle dans la synthèse de l' O-acétylserine.

Elle peut notamment être une SAT sensible à la cystéine, comme par exemple une SAT de plante, par exemple une SAT chloroplastique ou mitochondriale (SAT2, SAT4, SAT1), ou une SAT native d'origine bactérienne ([22] Nakamori & al., 1998,  
 10 Appl. Environ. Microbiol., 64, 1607-1611 ; [23] Takagi H. & al., 1999, Febs Lett. 452, 323-327 ; [24], Mino K. & al., 1999, Biosci. Biotechnol. Biochem., 63, 168-179).

Elle peut aussi être une SAT insensible à la cystéine, comme par exemple une SAT de plante, par exemple une SAT cytoplasmique (SAT3), ou une SAT d'origine bactérienne mutée, rendue insensible à la cystéine par mutagenèse ([22] et [23], dont les  
 15 contenus sont incorporés ici par référence), ou toute SAT de plantes mutée et fonctionnelle dans la synthèse de l'O-acétylsérine (le précurseur carboné pour la synthèse de la cystéine).

Selon un mode particulier de réalisation de l'invention, la SAT est une SAT d'*Arabidopsis thaliana* [12].

20 Selon un premier mode de réalisation de l'invention, la SAT est surexprimée dans le cytoplasme des cellules végétales. La SAT est soit une SAT cytoplasmique de plante, en particulier la SAT 3 (L34076) ou SAT3' ou SAT52 (U30298), représentée par la SEQ ID NO 1 ou la SEQ ID NO 2 respectivement, ou une SAT d'origine bactérienne telle que définie ci-dessus. La SAT surexprimée dans le cytoplasme peut également être  
 25 une SAT de plante non cytoplasmique, par exemple une SAT chloroplastique ou mitochondriale. Ces SAT non cytoplasmiques de plante sont naturellement exprimées dans le cytoplasme sous la forme d'une protéine précurseur comprenant un signal d'adressage vers le compartiment cellulaire différent du cytoplasme dans lequel la SAT mature fonctionnelle est libérée. Pour surexprimer ces SAT fonctionnelles matures dans  
 30 le cytoplasme, on les ampute de leur signal d'adressage. Dans ce cas, la protéine SAT surexprimée dans le cytoplasme est une SAT de plante non cytoplasmique amputée de son ou ses signaux d'adressage vers des compartiments cellulaires différents du cytoplasme.

Selon un mode particulier de réalisation de l'invention, la SAT non  
 35 cytoplasmique amputée de son signal d'adressage est la SAT1' représentée par la SEQ ID NO 3

Selon un deuxième mode de réalisation de l'invention, la SAT est surexprimée dans les mitochondries. La protéine est avantageusement exprimée dans le cytoplasme sous la forme d'une protéine de fusion peptide signal/SAT, la SAT mature fonctionnelle

étant libérée à l'intérieur des mitochondries. Le peptide signal d'adressage mitochondrial est avantageusement constitué par au moins un peptide signal d'adressage mitochondrial d'une protéine végétale à localisation mitochondriale, comme le peptide signal de la sous unité  $\beta$ -F1 ATPase de tabac [[25] Hemon P. & al. 1990, Plant Mol. Biol. 15, 895-904], ou le peptide signal de la SAT1 représenté par les acides aminés 1 à 63 sur la SEQ ID NO 4.

Selon un mode particulier de réalisation de l'invention, la SAT mitochondriale est la SAT1 (U22964) représentée par la SEQ ID NO 4.

Selon un troisième mode de réalisation de l'invention, la SAT est surexprimée dans les chloroplastes des cellules végétales.

La SAT sera exprimée dans les chloroplastes par tout moyen approprié, en particulier par tout moyen connu de l'homme du métier et largement décrit dans l'état de la technique.

Selon un mode particulier de réalisation de l'invention, la SAT est surexprimée dans les chloroplastes par intégration dans l'ADN chloroplastique d'un gène chimère comprenant une séquence d'ADN codant pour ladite SAT, sous le contrôle d'éléments de régulations en 5' et 3' fonctionnels dans les chloroplastes. Les techniques d'insertion de gènes dans les chloroplastes, comme les éléments de régulations appropriés à l'expression desdits gènes dans les chloroplastes sont bien connus de l'homme du métier, et notamment décrits dans les brevets et demandes de brevet suivants : US 5,693,507, US 5,451,513 et WO 97/32977.

Selon un autre mode particulier de réalisation de l'invention, la SAT est surexprimée dans le cytoplasme sous la forme d'une protéine de fusion peptide de transit/SAT, le peptide de transit ayant pour fonction d'adresser la SAT à laquelle il est fusionné vers les chloroplastes, la SAT mature fonctionnelle étant libérée à l'intérieur des chloroplastes après clivage au niveau de la membrane chloroplastique.

Dans ce cas, la SAT peut être une SAT chloroplastique d'origine végétale, comme la SAT2 ou la SAT4, représentée par la SEQ ID NO 5 ou 6 respectivement

La SAT peut également être une SAT cytoplasmique d'origine végétale ou une SAT d'origine bactérienne telles que définies précédemment. Dans les SAT cytoplasmiques on entend également les SAT non cytoplasmiques amputées de leur signal d'adressage vers un compartiment différent du cytoplasme telles que définies précédemment.

Les peptides de transit, leurs structures, leurs modes de fonctionnement et leur utilisation dans la construction de gènes chimères pour l'adressage d'une protéine hétérologue dans les chloroplastes, ainsi que des peptides de transit chimères comprenant plusieurs peptides de transit sont bien connus de l'homme du métier et largement décrits dans la littérature. On citera notamment les demandes de brevet

suivants : EP 189 707, EP 218 571 et EP 508 909, et les références citées dans ces demandes de brevet, dont le contenu est incorporé ici par référence

Dans la protéine de fusion selon l'invention, la SAT peut être homologue ou hétérologue du peptide de transit. Dans le premier cas, la protéine de fusion est la protéine SAT2 ou la SAT4 exprimée naturellement dans les chloroplastes des cellules végétales. Dans le second cas, le peptide de transit peut être un peptide de transit d'une SAT2, représentée par les acides aminés 1 à 32 de la SEQ ID 5, ou le peptide de transit d'une SAT4, représenté par les acides aminés 1 à 30 de la SEQ ID NO 6, ou encore un peptide de transit d'une autre protéine à localisation plastidiale, notamment les peptides de transit définis ci-après. On entend par protéine à localisation plastidiale une protéine exprimée dans le cytoplasme des cellules végétales sous la forme d'une protéine de fusion peptide de transit/protéine, la protéine mature étant localisée dans le chloroplaste après clivage du peptide de transit.

Un peptide de transit d'EPSPS de plante est notamment décrit dans la demande de brevet EP 218 571, alors qu'un peptide de transit de ssu RuBisCO de plante est décrit dans la demande de brevet EP 189 707.

Selon un autre mode de réalisation de l'invention, le peptide de transit comprend également, entre la partie C-terminale du peptide de transit et la partie N-terminale de la SAT une partie de séquence de la partie mature N-terminale d'une protéine à localisation plastidiale, cette partie de séquence comprenant généralement moins de 40 acides aminés de la partie N-terminale de la protéine mature, de préférence moins de 30 acides aminés, plus préférentiellement entre 15 et 25 acides aminés. Un tel peptide de transit comprenant un peptide de transit fusionné à une partie de la partie N-terminale d'une protéine à localisation plastidiale est notamment décrit dans la demande de brevet EP 189 707, plus particulièrement pour le peptide de transit et la partie N-terminale de ssu RuBisCO de plante.

Selon un autre mode de réalisation de l'invention, le peptide de transit comprend également entre la partie C-terminale de la partie N-terminale de la protéine mature et la partie N-terminale de la SAT un deuxième peptide de transit d'une protéine végétale à localisation plastidiale. Préférentiellement, ce peptide de transit chimère comprenant l'association de plusieurs peptide de transit est un peptide de transit optimisé (OTP) constitué par la fusion d'un premier peptide de transit, avec une partie de séquence de la partie mature N-terminale d'une protéine à localisation plastidiale, laquelle est fusionnée avec un deuxième peptide de transit. Un tel peptide de transit optimisé est décrit dans la demande de brevet EP 508 909, plus particulièrement le peptide de transit optimisé comprenant le peptide de transit de la ssu RuBisCO de tournesol, fusionné à un peptide constitué par les 22 acides aminés de la partie N-terminale de la ssu RuBisCO de maïs mature, fusionné au peptide de transit de la ssu RuBisCO de maïs.

La présente invention concerne également une protéine de fusion peptide de transit/SAT, dans laquelle la SAT définie plus haut est hétérologue du peptide de transit, et dans laquelle le peptide de transit est constitué par au moins un peptide de transit d'une protéine végétale naturelle à localisation plastidiale tel que défini ci-dessus.

- 5 La présente invention concerne également une séquence d'acide nucléique codant pour une protéine de fusion peptide de transit/SAT décrite ci-dessus. Selon la présente invention, on entend par "séquence d'acide nucléique" une séquence nucléotidique pouvant être de type ADN ou ARN, de préférence de type ADN, notamment double brin, qu'elle soit d'origine naturelle ou synthétique, notamment une
- 10 séquence d'ADN pour laquelle les codons codant pour la protéine de fusion selon l'invention auront été optimisés en fonction de l'organisme hôte dans lequel elle sera exprimée, ces méthodes d'optimisations étant bien connues de l'homme du métier.

- L'invention a encore pour objet l'utilisation d'une séquence d'acide nucléique codant pour une SAT selon l'invention définie précédemment, en particulier pour un
- 15 adressage cytoplasmique, mitochondrial ou chloroplastique, dans un procédé pour la transformation des plantes, comme séquence codante permettant de modifier le contenu en cystéine, méthionine et dérivés, et glutathion des plantes transformées. Cette séquence peut bien entendu également être utilisée en association avec d'autre(s) gène(s) marqueur(s) et/ou séquence(s) codante(s) pour une ou plusieurs autres propriétés
- 20 agronomiques.

- La présente invention concerne également un gène chimère (ou cassette d'expression) comprenant une séquence codante ainsi que des éléments de régulation en position 5' et 3' hétérologues pouvant fonctionner dans un organisme hôte, en particulier les cellules végétales ou les plantes, la séquence codante comprenant au moins une
- 25 séquence d'acide nucléique codant pour une SAT telle que définie précédemment.

- Par organisme hôte, on entend tout organisme mono ou pluricellulaire, inférieur ou supérieur, dans lequel le gène chimère selon l'invention peut être introduit. Il s'agit en particulier de bactéries, par exemple *E. coli*, de levures, en particulier des genres *Saccharomyces* ou *Kluyveromyces*, *Pichia*, de champignons, en particulier *Aspergillus*,
- 30 d'un baculovirus, ou de préférence des cellules végétales et des plantes.

Par "cellule végétale", on entend selon l'invention toute cellule issue d'une plante et pouvant constituer des tissus indifférenciés tels que des cals, des tissus différenciés tels que des embryons, des parties de plantes, des plantes ou des semences.

- On entend par "plante" selon l'invention, tout organisme multicellulaire différencié capable de photosynthèse, en particulier monocotylédones ou dicotylédones,
- 35 plus particulièrement des plantes de culture destinées ou non à l'alimentation animale ou humaine, comme le maïs, le blé, le colza, le soja, le riz, la canne à sucre, la betterave, le tabac, le coton, etc.

Les éléments de régulation nécessaires à l'expression de la séquence d'acide

nucléique codant pour une protéine de fusion selon l'invention sont bien connus de l'homme du métier en fonction de l'organisme hôte. Ils comprennent notamment des séquences promotrices, des activateurs de transcription, des séquences terminatrices, y compris des codons start et stop. Les moyens et méthodes pour identifier et sélectionner les éléments de régulation sont bien connus de l'homme du métier et largement décrits dans la littérature.

L'invention concerne plus particulièrement la transformation des plantes. Comme séquence de régulation promotrice dans les plantes, on peut utiliser toute séquence promotrice d'un gène s'exprimant naturellement dans les plantes en particulier un promoteur s'exprimant notamment dans les feuilles des plantes, comme par exemple des promoteurs dits constitutifs d'origine bactérienne, virale ou végétale ou encore des promoteurs dits lumière dépendants comme celui d'un gène de la petite sous-unité de ribulose- biscoxylase/oxygénase (RuBisCO) de plante ou tout promoteur convenable connu pouvant être utilisé. Parmi les promoteurs d'origine végétale on citera les promoteurs d'histone tels que décrits dans la demande EP 0 507.698, ou le promoteur d'actine de riz (US 5,641,876). Parmi les promoteurs d'un gène de virus de plante, on citera celui de la mosaïque du choux fleur (CAMV 19S ou 35S), ou le promoteur du circovirus (AU 689 311).

On peut encore utiliser une séquence de régulation promotrice spécifique de régions ou de tissus particuliers des plantes, et plus particulièrement des promoteurs spécifiques des graines ([26 ] Datla, R.& al., Biotechnology Ann. Rev. (1997) 3, 269-296), notamment les promoteurs de la napine (EP 255 378), de la phaseoline, de la glutenine, de la zéine, de l'héliantinine (WO 92/17580), de l'albumine (WO 98/45460), de l'oélosine (WO 98/45461), de l'ATS1 ou de l'ATS3 (WO 99/20275).

Selon l'invention, on peut également utiliser, en association avec la séquence de régulation promotrice, d'autres séquences de régulation, qui sont situées entre le promoteur et la séquence codante, telles que des activateurs de transcription ("enhancer"), comme par exemple l'activateur de translation du virus de la mosaïque du tabac (TMV) décrit dans la demande WO 87/07644, ou du virus etch du tabac (TEV) décrit par Carrington & Freed.

Comme séquence de régulation terminatrice ou de polyadénylation, on peut utiliser toute séquence correspondante d'origine bactérienne, comme par exemple le terminateur nos d'*Agrobacterium tumefaciens*, ou encore d'origine végétale, comme par exemple un terminateur d'histone tel que décrit dans la demande EP 0 633 317.

La présente invention concerne également un vecteur de clonage et/ou d'expression pour la transformation d'un organisme hôte contenant au moins un gène chimère tel que défini ci-dessus. Ce vecteur comprend outre le gène chimère ci-dessus, au moins une origine de répllication. Ce vecteur peut être constitué par un plasmide, un cosmide, un bactériophage ou un virus, transformés par l'introduction du gène chimère

selon l'invention. De tels vecteurs de transformation en fonction de l'organisme hôte à transformer sont bien connus de l'homme du métier et largement décrits dans la littérature. Pour la transformation des cellules végétales ou des plantes, il s'agira notamment d'un virus qui peut être employé pour la transformation des plantes  
 5 développées et contenant en outre ses propres éléments de réplication et d'expression. De manière préférentielle, le vecteur de transformation des cellules végétales ou des plantes selon l'invention est un plasmide.

Pour la transformation des organismes hôtes, le gène chimère selon l'invention peut être employé en association avec un gène marqueur de sélection, soit dans un  
 10 même vecteur, les deux gènes étant associés de manière convergente, divergente ou colinéaire, ou encore dans deux vecteurs employés simultanément pour la transformation de l'organisme hôte. De tels gènes marqueurs et leur utilisation pour la transformation d'organismes hôtes sont bien connus de l'homme du métier et largement décrits dans la littérature.

Parmi les gènes codant pour des marqueurs de sélection, on peut citer les gènes de résistance aux antibiotiques, les gènes de tolérance aux herbicides (bialaphos, glyphosate ou isoxazoles), des gènes codant pour des enzymes facilement identifiables comme l'enzyme GUS (ou GFP, "Green Fluorescent Protein"), des gènes codant pour des pigments ou des enzymes régulant la production de pigments dans les cellules  
 20 transformées. De tels gènes marqueurs de sélection sont notamment décrits dans les demandes de brevet EP 242 236, EP 242 246, GB 2 197 653, WO 91/02071, WO 95/06128, WO 96/38567 ou WO 97/04103.

L'invention a encore pour objet un procédé de transformation des organismes hôtes, en particulier des cellules végétales par intégration d'au moins une séquence  
 25 d'acide nucléique ou un gène chimère tels que définis ci-dessus, transformation qui peut être obtenue par tout moyen connu approprié, amplement décrit dans la littérature spécialisée et notamment les références citées dans la présente demande, plus particulièrement par le vecteur selon l'invention.

Une série de méthodes consiste à bombarder des cellules, des protoplastes ou des  
 30 tissus avec des particules auxquelles sont accrochées les séquences d'ADN. Une autre série de méthodes consiste à utiliser comme moyen de transfert dans la plante un gène chimère inséré dans un plasmide Ti d'*Agrobacterium tumefaciens* ou Ri d'*Agrobacterium rhizogenes*. D'autres méthodes peuvent être utilisées telles que la micro-injection ou l'électroporation, ou encore la précipitation directe au moyen de  
 35 PEG. L'homme du métier fera le choix de la méthode appropriée en fonction de la nature de l'organisme hôte, en particulier de la cellule végétale ou de la plante.

La présente invention a encore pour objet les organismes hôtes, en particulier cellules végétales ou plantes, transformés et contenant un gène chimère défini ci-dessus.

La présente invention a encore pour objet les plantes contenant des cellules

transformées, en particulier les plantes régénérées à partir des cellules transformées. La régénération est obtenue par tout procédé approprié qui dépend de la nature de l'espèce, comme par exemple décrit dans les références ci-dessus. Pour les procédés de transformation des cellules végétales et de régénération des plantes, on citera  
 5 notamment les brevets et demandes de brevet suivants: US 4,459,355, US 4,536,475, US 5,464,763, US 5,177,010, US 5,187,073, EP 267.159, EP 604 662, EP 672 752, US 4,945,050, US 5,036,006, US 5,100,792, US 5,371,014, US 5,478,744, US 5,179,022, US 5,565,346, US 5,484,956, US 5,508,468, US 5,538,877, US 5,554,798, US 5,489,520, US 5,510,318, US 5,204,253, US 5,405,765, EP 442 174, EP 486 233, EP  
 10 486 234, EP 539 563, EP 674 725, WO 91/02071 et WO 95/06128.

La présente invention concerne également les plantes transformées issues de la culture et/ou du croisement des plantes régénérées ci-dessus, ainsi que les graines de plantes transformées.

Les plantes transformées pouvant être obtenues selon l'invention peuvent être du  
 15 type monocotylédones telles que par exemple les céréales, la canne à sucre, le riz et le maïs ou dicotylédones comme par exemple le tabac, la soja, le colza, le coton, la betterave, le trèfle, etc.....

Les plantes transformées selon l'invention peuvent contenir d'autres gènes d'intérêt, conférant aux plantes de nouvelles propriétés agronomiques. Parmi les gènes  
 20 conférant de nouvelles propriétés agronomiques aux plantes transformées, on peut citer les gènes conférant une tolérance à certains herbicides, ceux conférant une tolérance à certains insectes, ceux conférant une tolérance à certaines maladies . De tels gènes sont notamment décrits dans les demandes de brevet WO 91/02071 et WO 95/06128. On peut également citer les gènes modifiant la constitution des plantes modifiées, en  
 25 particulier la teneur et la qualité de certains acides gras essentiels (EP 666 918) ou encore la teneur et la qualité des protéines, en particuliers dans les feuilles et/ou les graines desdites plantes. On citera en particulier les gènes codant pour des protéines enrichies en acides aminés soufrés ([1]; WO 98/20133 ; WO 97/41239 ; WO 95/31554 ; WO 94/20828 ; WO 92/14822, US 5 939 599, US 5 912 424). Ces protéines enrichies  
 30 en acides aminés soufrés auront également pour fonction de piéger et stocker la cystéine et/ou la méthionine excédentaire, permettant d'éviter les problèmes éventuels de toxicité liés à une surproduction de ces acides aminés soufrés en les piégeant.

On peut citer également des gènes codant pour des peptides riches en acides aminés soufrés et plus particulièrement en cystéines, les dits peptides ayant également  
 35 une activité antibactérienne et/ou antifongique. On citera plus particulièrement les défensines de plantes, de même que les peptides lytiques de toute origine, et plus particulièrement les peptides lytiques suivants : l'androctonine (WO 97/30082 et WO 99/09189), la drosomicine (WO 99/02717), la thanatine (WO 99/24594) ou l'héliomicine (WO 99/53053).

Ces autres gènes d'intérêt peuvent être combinés au gène chimère selon l'invention soit par croisement conventionnel de deux plantes contenant chacune l'un des gènes (la première le gène chimère selon l'invention et la seconde le gène codant pour la protéine d'intérêt) soit en effectuant la transformation de cellules végétales d'une plante contenant le gène codant pour la protéine d'intérêt avec le gène chimère selon l'invention.

Les exemples ci-après permettent d'illustrer l'invention, sans toutefois chercher à en limiter la portée.

Toutes les méthodes ou opérations décrites ci-dessous dans ces exemples sont données à titre d'exemples et correspondent à un choix, effectué parmi les différentes méthodes disponibles pour parvenir au même résultat. Ce choix n'a aucune incidence sur la qualité du résultat et par conséquent, toute méthode adaptée peut être utilisée par l'homme de l'art pour parvenir au même résultat. La plupart des méthodes d'ingénierie des fragments d'ADN sont décrites dans "Current Protocols in Molecular Biology" Volumes 1 et 2, Ausubel F.M. et al, publiés par Greene Publishing Associates et Wiley-Interscience (1989) ou dans Molecular cloning, T.Maniatis, E.F.Fritsch, J.Sambrook, 1982.

Le contenu de toutes les références citées dans la description ci-dessus et dans les exemples ci-après est incorporé au contenu de la présente demande de brevet par référence.

#### **Exemple 1. Mise en évidence de l'inhibition de la serine acetyltransférase chloroplastique de feuilles de pois (*Pisum sativum*) par la cysteine**

On prépare à partir de feuilles de pois les trois compartiments subcellulaires correspondant au cytosol (préparation à partir d'un fractionnement subcellulaire de protoplastes de pois, [12]), aux mitochondries et aux chloroplastes [12]. On en extrait les protéines solubles et on mesure l'activité sérine acétyltransférase présente dans chacun des compartiments au moyen d'une technique décrite [12, 17].

#### **Description de la méthode de dosage:**

L'activité de la serine acetyltransférase est mesurée par chromatographie liquide haute performance (HPLC), en dosant l'O-acétylserine formée au cours de la réaction (réaction 1), après dérivatization avec de l'orthophtalaldéhyde (OPA). Une quantité définie de l'extrait protéique, correspondant au cytosol, et aux différentes fractions solubles des chloroplastes (stroma) et mitochondries (matrice), est dessalé sur une colonne PD10 (Pharmacia) préalablement équilibrée en tampon 50 mM Hepes-HCl, pH 7,5 et 1 mM EDTA. La réaction enzymatique est effectuée en présence de 50 mM Hepes-HCl, pH 7.5, 1mM dithiotreitol, 10 mM L-serine, 2.5 mM acétyl-CoA, dans un volume réactionnel de 100 µl, à 25 °C. Après 10 à 15 min d'incubation, la réaction est

- arrêtée par l'addition de 50 µl d'acide trichloroacétique 20% (P/V). Les protéines ainsi précipitées sont ensuite éliminées par une centrifugation de 2 min à 15,000g. Le surnageant, contenant le produit de la réaction (OAS), est mélangé avec 500 µl d'une solution de dérivatization (54 mg d'OPA dissous dans 1 ml d'éthanol absolu, 9 ml d'une solution de borate-NaOH 400 mM, pH 9,5 et 0,2 ml de β-mercaptoethanol 14 M) et incubé pendant 2 min. Une fraction de ce mélange (20 µl) est injectée sur une colonne de phase inverse (colonne AccQ Tag C<sub>18</sub>, 3,9 X 150 mm, Waters) connectée à un système HPLC. Les tampons utilisés pour l'élution des composés dérivés par l'OPA sont: Tampon A, 85 mM acétate de sodium, pH 4.5, et acétonitrile 6% (V/V), pH 4,5; 10 Tampon B, acétonitrile, 60 % (V/V) dans de l'eau. L'*O*-acétylsérine, dérivée par l'OPA, est éluée par un gradient linéaire continu du tampon B dans le tampon A de 25 à 70 % (V/V), et détectée par fluorescence à 455 nm (excitation à 340 nm). Le temps de rétention de l'*O*-acétylsérine est dans nos conditions de l'ordre de 6.2 min, et la quantité de produit formé dans les essais enzymatiques est quantifiée à partir d'une courbe étalon 15 réalisée avec de l'*O*-acétylsérine. Les essais enzymatiques ont été optimisés afin de respecter le pH optimal de la réaction, la linéarité en fonction du temps, et afin d'opérer dans des concentrations saturantes en substrats.

#### Effet de la cystéine sur l'activité sérine acétyltransférase de feuilles de pois

- 20 L'incubation (2 min) est réalisée en présence de l'extrait protéique (cytosol, matrice et stroma), et des concentrations croissantes de L-cystéine (de 0 à 1 mM), avant l'ajout de concentrations saturantes des substrats de la sérine acétyltransférase, la L-sérine (10 mM) et l'acétyl-CoA (2,5 mM). La réaction enzymatique et le dosage de l'*O*-acétylsérine résiduelle dans le surnageant sont effectués comme décrit précédemment.
- 25 Le résultat de ces expériences est représenté sur le graphe de la **figure 2** en annexe.
- Si on ajoute de la cystéine libre (de 0 à 1 mM, **Figure 2**) aux différents essais, on constate une très forte inhibition de l'activité sérine acétyltransférase chloroplastique (constante d'inhibition de l'ordre de 30 µM). L'activité sérine acétyltransférase mitochondriale est inhibée mais pour des concentrations supérieures de cystéine 30 (constante d'inhibition de l'ordre de 300 µM). En revanche, l'activité sérine acétyltransférase cytosolique est insensible à l'inhibition par la cystéine jusqu'à des concentrations de l'ordre de 1 mM (**figure 2**). Ce résultat prouve donc que seule l'activité sérine acétyltransférase chloroplastique, donc l'enzyme associée à la voie de l'assimilation du sulfate, est inhibée par le produit final, la L-cystéine.

**Tableau I: Détermination des activités spécifiques et des valeurs de CI<sub>50</sub> pour la cystéine pour chacune des isoformes de la serine acétyltransférase.**

Serine acétyltransférase ( <i>Pisum sativum</i> )		
	Activité spécifique nmoles OAS • min <sup>-1</sup> • mg <sup>-1</sup>	CI <sub>50</sub> L-cystéine μM
<b>Stroma</b>	0,93 ± 0,2	33,4 ± 8
<b>Matrice</b>	10 ± 2	283 ± 50
<b>Cytosol</b>	0,83 ± 0,3	pas d'inhibition

- La concentration en L-cystéine permettant l'obtention de 50% d'inhibition (CI<sub>50</sub>) dans les conditions standards de la réaction calculée pour différentes préparations enzymatiques est représentée dans le tableau I. Les déterminations de l'activité enzymatique de la serine acétyltransférase et de la CI<sub>50</sub> ont été effectuées lors de 9 expériences différentes (stroma), et lors de 3 pour les extraits cytosolique et issus de la mitochondrie. De même, l'activité de la serine acétyltransférase du chloroplaste de feuilles d'épinards est sensible à la cystéine. A l'opposée, chez *Arabidopsis thaliana* il apparaît que seul l'activité de l'isoforme associée au compartiment cytosolique soit contrôlée par le niveau de cystéine ([27] Noji M. & al. 1998, J. Biol. Chem. 273, 32739-32745 ; [28] Inoue K. & al. 1999, Eur. J. Biochem. 266, 220-227). Pour ces auteurs, l'activité associée au compartiment chloroplastique est insensible à la cystéine. Il semblerait donc que l'inhibition de l'activité serine acétyltransférase chloroplastique par la cystéine soit un phénomène plante-spécifique, mais en particulier très marquée chez les légumineuses comme le pois.

#### Etude du mode d'inhibition de l'activité serine acétyltransférase par la cystéine

- La vitesse de la réaction enzymatique a été déterminée pour des concentrations fixes de cystéine (0 μM; 10 μM; 20 μM; 40 μM; 60 μM et 100 μM) en faisant varier soit la concentration en L-serine soit en acétyl-CoA, pour des concentrations saturantes du second co-substrat. Pour chacune des cinétiques obtenues, l'affinité de l'enzyme pour ces substrats ne semble pas être affectée, mais par contre la vitesse maximale de la réaction est modifiée. Plus la concentration en L-cystéine augmente, plus la vitesse de synthèse de l'O-acétylserine est diminuée. Pour chacune des conditions analysées, la constante d'inhibition  $K_i$  a été estimée de l'ordre de 30 (± 2,2) μM (substrat variable: la serine), et de 22 (± 2 μM) (substrat variable: l'acétyl-CoA). Nous avons pu montrer que la cystéine est un inhibiteur de type non-compétitif pour l'activité serine acétyltransférase et de plus de type allostérique (constante de Hill de l'ordre de 1,6 ± 0,3 μM) en utilisant les équations cinétiques classiques ([29] Segel, I.H. (1995), John Wiley and Sons, New-York). Ces résultats indiquent que l'inhibition de l'enzyme

chloroplastique prend place en un site différent du site actif, et de plus qui n'existe pas sur l'isoforme serine acétyltransférase, présente dans le cytosol.

Ces valeurs de constantes d'inhibition sont consistantes avec la concentration en cystéine déterminée dans les chloroplastes de pois de  $40 \pm 10 \mu\text{M}$  (2 nmoles / mg chlorophylle), valeur qui est calculée pour un volume du compartiment stromatique de l'ordre de 35 à 65  $\mu\text{l}$  par mg de chlorophylle.

#### Dissociation du complexe bi-enzymatique, cystéine synthase, par la cystéine

La serine acétyltransférase de la cellule végétale, comme son homologue bactérien, forme un complexe enzymatique avec l'*O*-acétylserine (thiol) lyase, l'enzyme qui catalyse la condensation du soufre réduit avec l'*O*-acétylserine. Ce complexe bi-enzymatique est appelée cystéine synthase. Toute la serine acétyltransférase du chloroplaste existe sous une forme en complexe avec l'*O*-acétylserine (thiol) lyase, alors que la majorité de l'*O*-acétylserine (thiol) lyase est sous la forme libre. La distribution de chacune de ces enzymes dans chacun des compartiments sub-cellulaires de feuilles de pois est décrite dans le Tableau II.

**Tableau II : Activité spécifique des activités serine acétyltransférase et *O*-acétylserine (thiol) lyase des compartiments cellulaires de feuilles de pois**

	Serine acétyltransférase	<i>O</i> -acétylserine (thiol) lyase	Ratio OASTL / SAT
	Activité spécifique (mU/mg)		
<b>Stroma</b>	0.85	260	306
<b>Matrice</b>	12	50	4
<b>Cytosol</b>	0.90	180	200

Le rapport de l'activité *O*-acétylserine (thiol) lyase (OASTL) à l'activité serine acétyltransférase (SAT) rend compte du large excès de l'OASTL sur la SAT. En particulier dans le stroma (chloroplaste) où s'effectue l'assimilation et la réduction du sulfate, et dans le cytosol, 95% de l'activité OASTL est sous forme libre. Ces conditions sont nécessaires pour une synthèse optimale de la cystéine [14]. Le complexe cystéine synthase est composée d'un tétramère de serine acétyltransférase et de deux dimères d'*O*-acétylserine (thiol) lyase. L'*O*-acétylserine, le produit de la réaction de la serine acétyltransférase, dissocie ce complexe bienzymatique, et le soufre tend à stabiliser celui-ci [14]. Ces interactions protéines-protéines au sein du complexe confèrent de nouvelles propriétés à chacune des enzymes, en particulier la serine acétyltransférase acquière de nouvelles propriétés catalytiques comparativement à la forme libre. De plus,

l'*O*-acétylserine (thiol) lyase en complexe est inactive dans la synthèse de la cystéine, et seul la forme libre (en excès dans la cellule) catalyse la synthèse de la cystéine [14].

Une fraction chloroplastique (*Pisum sativum*), préalablement incubée en présence d'une concentration optimale de cystéine (0.1 mM), conduisant à l'inhibition de la serine acétyltransférase (voir figure 2) est ensuite soumise à une chromatographie de gel filtration permettant la séparation des molécules en fonction de leur masse moléculaire. Dans ces conditions le complexe cystéine synthase se dissocie en tétramères de serine acétyltransférase et dimères d'*O*-acétylserine (thiol) lyase. La serine acétyltransférase chloroplastique sous sa forme libre est toujours sensible à l'inhibition par la cystéine. Pour affiner ce résultat et confirmer que l'inhibition de l'enzyme n'est pas dépendante de l'interaction avec l'OASTL, une serine acétyltransférase a été partiellement purifiée à partir de chloroplaste de pois, par une chromatographie d'échanges d'ions suivie par une chromatographie de filtration moléculaire réalisée en présence d'*O*-acétylserine (1 mM), une condition qui entraîne la dissociation du complexe.

La fraction de serine acétyltransférase ainsi libre de contaminations par l'*O*-acétylserine (thiol) lyase est incubée en présence de concentrations croissantes de cystéine dans les conditions décrites dans le tableau I et Figure 2. La  $CI_{50}$  calculée est de l'ordre de 15 +/- 3 micromolaire et est comparable à la valeur obtenue précédemment pour l'enzyme dans les conditions du chloroplaste (voir Tableau I). Ce dernier résultat permet d'établir un modèle pour rendre compte de l'inhibition de la serine acétyltransférase chloroplastique. Dans la figure 3, la forme tétramérique de la serine acétyltransférase (SAT) est figurée par les cercles gris et le dimère de l'*O*-acétylserine (thiol) lyase (OASTL) par les cercles noirs. Le complexe cystéine synthase fonctionnelle dans la cellule est figuré par l'association des deux populations moléculaires. En présence de cystéine, le complexe cystéine synthase fixe de la cystéine qui modifie les interactions protéines-protéines au sein du complexe cystéine synthase, et conduit à la dissociation en tétramères de SAT et dimères d'OASTL. La SAT ainsi sous sa forme libre est aussi sensible à la cystéine, et l'on sait que cette structure a tendance à former des agrégats (hors du complexe cystéine synthase) et qui se traduisent par une perte de son activité [14].

#### **Exemple 2. Isolement et caractérisation d'un gène codant pour une isoforme serine acétyltransférase putative cytoplasmique (SAT3) [12]**

On reprend dans cet exemple le mode opératoire décrit en page 502 de Ruffet & al. [12], en particulier les chapitres décrits sous les titres " Bacterial strain and growth conditions " et " Isolation of *A. thaliana* serine acetyltransferase cDNA clones by complementation in *E. coli* ".

Un gène codant pour une serine acétyltransférase putative cytosolique (Z34888 ou L34076) représentée sur la **Figure 4** (SEQ ID NO 1), a été isolé par complémentation fonctionnelle d'une souche d'*Escherichia coli* déficiente pour l'activité serine acétyltransférase. L'analyse de la séquence primaire a montré la présence d'une forte similitude avec la séquence de l'enzyme bactérienne (56% d'homologie et 41% d'identité).

Les amorces qui ont été utilisées pour l'amplification de la séquence nucléotidique et son clonage dans le vecteur utilisé pour la transformation de plants de tabac sont les suivantes :

10 Oligo 1 : 5' GAGAGAGGAT CCTCTTTCCA ATCATAAACC ATGGCAACAT  
GCATAGACAC ATGC 3'  
Oligo 2 : 5' GGCTCACCAG ACTAATACAC TAAATTGTGT TTACCCTCGAG  
AGAGAG 3'

Ces amorces permettent l'introduction des sites de restriction *Bam*HI en 5' (GGATCC) et *Sac*I en 3' (GAGCTC).

L'extrémité N-terminale de la séquence en acide aminé de l'isoforme SAT3 ne présente pas les caractéristiques des peptides d'adressage dans un organe (mitochondrie ou chloroplaste). Cette analyse conduit à supposer une localisation cytosolique pour cette isoforme [12]. L'absence de peptide d'adressage de type chloroplastique pour cette isoforme a pu être confirmé lors d'expérience d'import dans les chloroplastes ([29] Murillo & al. 1995, Cell. and Mol. Biol. Research 41, 425-433). A l'opposé, une étude utilisant des constructions incluant une portion en de la séquence nucléotidique et une protéine marqueur (Green Fluorescent Protein, GFP) ont montré que la présence du produit de fusion (5'-SAT3-GFP) dans le chloroplaste de plants d' *A. thaliana* transformés (stade végétatif de la plante) et aussi dans le cytosol (au stade florale) [27].

Le gène de la SAT 3 (L34076) présente une structure sans introns.

### Exemple 3. Sur-expression et purification de la SAT3 chez *Escherichia coli*

Le protocole défini pour la sur-expression de l'enzyme chez *E. coli* permet la purification de l'enzyme sous sa forme libre ou en complexe avec l'*O*-acétylserine (thiol) lyase de plante, le complexe cystéine synthase [14]. Avec les protéines purifiées, l'effet de la cystéine sur l'activité de la serine acétyltransférase a été analysée par un dosage spectrophotométrique basé sur la consommation de l'acétyl-CoA au cours de la réaction 1, en fonction du temps d'incubation. Cette analyse s'effectue dans un milieu (1 ml) contenant 50 mM Hepes-HCl, pH 7.5, 2 mM L-serine et 0.2 mM Acetyl-CoA. La réaction est suivie en mesurant la diminution de l'absorbance à 232 nm (coefficient d'extinction moléculaire de 4200 M<sup>-1</sup>cm<sup>-1</sup>) ([30] Kredich, N.M. & al., J. Biol. Chem. (1969) 244, 2428-2439). Nous avons pu montrer que cette isoforme (SAT3) sous sa

forme libre ou en complexe avec l'O-acétylserine (thiol) lyase est insensible à la cystéine. Ce résultat nous permet de confirmer que cet ADNc (L34076, **Figure 4**) code pour une serine acétyltransférase cytosolique, car la composition en acides aminés de l'extrémité N-terminale ne présente pas les caractéristiques de peptides de transit, et de plus cette sérine acétyltransférase est insensible à la cystéine. Ce dernier résultat est similaire aux observations obtenues pour l'activité serine acétyltransférase cytosolique de feuilles de pois (**Figure 2** et Tableau I).

**Exemple 4. Isolement et caractérisation d'un gène codant pour une isoforme serine acétyltransférase cytoplasmique (SAT3') (U30298)**

On reprend le mode opératoire de l'exemple 3 avec les oligonucléotides 3 et 4 suivants :

Oligo 3 : 5' GAGAGAGGAT CCTCTTATCG CCGCGTTAAT ATGCCACCGG  
CCGGAGAACTC C 3'

Oligo 4 : 5' GAGCCTTACC AGTCTAATGT AGTATATTTC AACCTCGAGA  
GAGAG 3'

On isole un gène codant pour une acétyltransférase (U 30298) représentée sur la **figure 5** (SEQ ID NO 2). L'analyse de la séquence primaire a montré la présence d'une forte similitude avec la séquence de l'enzyme bactérienne (51.6 % d'homologie et 42.6 % d'identité). La structure du N-terminale (absence des conditions nécessaires permettant un adressage dans un organites) indique que cette isoforme possède une localisation cytosolique. Par contre elle est donnée sensible à la cystéine [27]. Ce résultat diffère des données obtenues avec les feuilles de pois (et d'épinards), dans le sens que le site de régulation par la cystéine semble être confinée au cytosol chez *A. thaliana*. [27]. De plus, il semblerait que *A. thaliana* possède au moins deux isoformes cytosoliques : SAT3 (exemple 3) et SAT3' (ou U30298, exemple 4). A la différence du gène de la SAT3, le gène correspondant à la SAT3' présente un intron.

**Exemple 5. Isolement et caractérisation de gènes codant pour une isoforme serine acétyltransférase (SAT1')**

Le mode opératoire décrit pour l'exemple 3 est repris pour le présent exemple.

Un gène codant pour une serine acétyltransférase (L78443) représenté sur la **figure 6** (SEQ ID NO 3) a été isolée par complémentation fonctionnelle d'une souche d'*Escherichia coli* déficiente pour l'activité serine acétyltransférase.[12] L'analyse de la séquence primaire montre de fortes similitudes avec la séquence de l'enzyme bactérienne (52.7% d'homologie et 39% d'identité).

Les amorces qui ont été utilisées pour l'amplification de la séquence nucléotidique et son clonage dans le vecteur utilisé pour la transformation de plants de tabac (en caractères gras sur la **figure 3**) sont les suivantes :

Oligo 5 : 5' GAGAGAGGAT CCCCTCCTCC TCCTCCTCCT ATGGCTGCGT  
 5 GCATCGACAC CTG 3'  
 Oligo 6 : 5' GCTCACCAGC CTAATACATT AAACCTTTTTC AGCTCGAGAG  
 AGAG 3'

Ces amorces permettent l'introduction des sites de restriction *Bam*HI en 5' (GGATCC) et *Sac*I en 3' (GAGCTC).

10 On obtient un deuxième gène codant pour une serine acétyltransférase putative mitochondriale (U22964) représenté sur la **figure 7** (SEQ ID NO 4) en reprenant le même mode opératoire avec l'oligo 7 en remplacement de l'oligo 5 comme amorce en 5'.

Oligo 7° : 5' GAGAGAGGAT CCGGCCGAGA AAAAAAAAAA ATGTTGCCGG  
 15 TCACAAGTCG CCG 3'

L'extrémité N-terminale de la séquence en acide aminé de l'isoforme SAT1 présente les caractéristiques des peptides d'adressage dans un organe (mitochondrie ou chloroplaste). Une localisation mitochondriale a été confirmée récemment par la  
 20 construction d'une protéine de fusion incluant la portion 5' et la "green fluorescent protéine" (5'SAT1-GFP) et par transformation de plants d'*Arabidopsis thaliana* [27]. Le gène de la SAT1' (L78443) ou SAT1 (U22964), comme son homologue (SAT3) ne présente pas d'intron.

#### 25 **Exemple 6. Sur-expression et purification de la SAT1 chez *Escherichia coli*. Localisation de cette isoforme chez *A. thaliana***

Le protocole défini pour la sur-expression de l'enzyme chez *E. coli* permet la purification de l'enzyme (sous sa forme sans peptide de transit, SAT L78443) sous sa forme libre ou en complexe avec l'*O*-acétylserine (thiol) lyase de plante, le complexe  
 30 cystéine synthase [14]. Avec les protéines purifiées, l'effet de la cystéine sur l'activité de la serine acétyltransférase a été analysée par un dosage spectrophotométrique basé sur la consommation de l'acétyl-CoA au cours de la réaction 1, en fonction du temps d'incubation (voir exemple 3). L'analyse a été aussi effectuée par un dosage du produit de la réaction (OAS) par HPLC (voir exemple 1). Nous avons pu montrer que cette  
 35 isoforme (SAT1') sous sa forme libre ou en complexe avec l'*O*-acétylserine (thiol) lyase est insensible à la cystéine. Ce dernier résultat parallèle les observations obtenues pour l'activité serine acétyltransférase mitochondriale de feuilles de pois (**Figure 2** et Tableau I). Cette dernière étant inhibée pour des concentrations non-physiologiques de cystéine.

Avec une préparation de mitochondries obtenue à partir de feuilles de pois, ou à partir de protoplastes issues de cultures de cellules, la localisation de cette isoforme dans la mitochondrie a pu être confirmée.

La fraction mitochondriale dépourvue de contaminations plastidiale et cytosolique a été obtenue en utilisant le protocole défini pour les mitochondries de feuilles de pois [12]. La masse moléculaire du polypeptide révélé par les anticorps dirigés contre le peptide [-TKTLHTRPILLEDLR-] (voir séquence acide aminé de la SAT1) est de l'ordre de 34000 daltons, une valeur qui est en accord avec la masse de la protéine obtenue grâce aux programmes d'analyse de séquence pour la prédiction des sites de clivages.

#### **Exemple 7. Isolement et caractérisation de gènes codant pour une isoforme serine acétyltransférase (SAT2)**

Le mode opératoire décrit pour l'exemple 3 est repris pour le présent exemple.

Un gène codant pour une serine acétyltransférase (L78444) représenté sur la **figure 8** (SEQ ID NO 5) a été isolé par complémentation fonctionnelle d'une souche d'*Escherichia coli* déficiente pour l'activité serine acétyltransférase.[12] L'analyse de la séquence primaire a montré la présence de fortes similitudes avec la séquence de l'enzyme bactérienne (49.5% d'homologie et 35.4% d'identité).

Les amorces qui ont été utilisées pour l'amplification de la séquence nucléotidique et son clonage dans le vecteur utilisé pour la transformation de plants de tabac (en caractères gras sur la **figure 8** sont les suivantes :

Oligo 8 : 5' **GAGAGAGGAT** CCGACAAGTT GGCATAATTT  
ATGGTGGATC TATCTTCCT 3'

Oligo 9 : 5' **CCTGTGTGAT** TGTCGTGTAG TACTCTAGAA  
ACTCGAGAGA GAG 3'

Ces amorces permettent l'introduction des sites de restriction *Bam*HI en 5' (GGATCC) et *Sac*I en 3' (GAGCTC).

L'analyse de la portion N-terminale de la séquence présente des caractéristiques pour un adressage de la protéine dans un organe (mitochondrie ou chloroplaste). A la différence des autres isoformes décrites ci-dessus, le gène de la SAT 2 est complexe et présente plusieurs introns. La comparaison des séquences de la SAT2 avec ses homologues d'*A. thaliana*, de plantes, et d'autres organismes laissent supposer une origine de type procaryotique (**Figure 10**). De plus, l'analyse de la séquence N-terminale en utilisant le programme chloroP [<http://www.cbs.dtu.dk/services/chlorP/>], indique de forte probabilité pour la présence d'un peptide de transit de type chloroplastique.

**Exemple 8. Isolement et caractérisation de gènes codant pour une isoforme serine acétyltransférase (SAT4)**

Un gène codant pour une serine acétyltransférase (SAT4) représenté sur la **figure 9** (SEQ ID NO 6) a été isolée par complémentation fonctionnelle d'une souche d'*Escherichia coli* déficiente pour l'activité serine acétyltransférase.[12] L'analyse de la séquence primaire a montré la présence de fortes similitudes avec la séquence de l'enzyme bactérienne (44.5% d'homologie et 32% d'identité)

Les amorces qui ont été utilisées pour l'amplification de la séquence nucléotidique et son clonage dans le vecteur utilisé pour la transformation de plants de tabac sont les suivantes :

Oligo 10 : 5' **GAGAGAGGAT CCGACAAGTTGG CATAATTTAT GGCTTGTATA AACGGCGAGA ATCGTGATTT TTCTT** 3'

Oligo 11: 5' **TACCTCGTAC CACTCAGAAC TCTAGAACT CGAGAGAGAG3'**

Ces amorces permettent l'introduction des sites de restriction *Bam*HI en 5' (GGATCC) et *Sac*I en 3' (GAGCTC).

L'analyse de la portion N-terminale de la séquence présente des caractéristiques pour un adressage de la protéine dans un organe (mitochondrie ou chloroplaste). Le gène de la SAT 4, comme celui de la SAT2, est complexe et présente plusieurs introns. La comparaison des séquences de la SAT4 avec ses homologues d'*A. thaliana*, de plantes, et autres organismes laissent supposer une origine de type procaryotique (**Figure 10**). De plus, l'analyse de la séquence N-terminale en utilisant le programme ChloroP [<http://www.cbs.dtu.dk/services/chlorP/>], indique une forte probabilité pour la présence d'un peptide de transit de type chloroplastique. La **figure 10** représente la comparaison des séquences, elle a été réalisée en utilisant le programme Clustaw (Vector NTI software). La SAT2 et la SAT4 sont plus proches des SATs procaryotiques que ne le sont les SAT3, SAT1 et SAT52. De plus, l'embranchement comprend aussi une SAT d'algue rouge (AB00848) identifié comme une protéine possédant une localisation chloroplastique et sensible à la cystéine ([32] Toda & al. 1998, Biochim. Biophys. Acta 1403, 72-84). La SAT4 est identifiée sur le chromosome 4 (Bac clone F8D20, accession AL031135).

**Exemple 8. Constructions utilisées pour la transformation des plants de tabac variété petit Havana**

**Expression du transgène dans les feuilles**

Les transformations des plants de tabac sont réalisées par l'intermédiaire d'*Agrobacterium tumefaciens* EHA105, contenant le vecteur pBI121 (Clonetch) (**Figures 11 et 12**).

### SAT3 (ou SAT1' ou toute SAT insensible à la cystéine)

Afin d'obtenir une expression de la SAT3 (SEQ ID NO 1) de l'exemple 2 dans le chloroplaste (Figure 11), on introduira en position 5' de l'ADNc une extension qui permettra l'adressage dans ce compartiment. Pour cela, le peptide de transit optimisé décrit auparavant sera utilisé.

Entre les bordures gauche (BG) et droite (BD) du T-DNA est cloné un gène de résistance à la kanamycine (NPTII) codant pour la néomycine phosphotransférase utilisé comme marqueur de sélection pour la transformation du tabac. L'expression du gène NPTII est sous la dépendance du promoteur et du terminateur de la nopaline synthase de *A. tumefaciens*. En aval, le gène de la  $\beta$ -glucuronidase cloné entre les sites uniques *Bam*H1 et *Sac*I, est sous le contrôle du promoteur 35S du virus de la mosaïque du chou-fleur (CaMV) et le signal de polyadénylation du gène de la nopaline synthase du plasmide Ti. L'insertion du produit de la construction OTP-SAT3 s'effectue entre les sites *Xho* et *Sac*I du vecteur délété du gène de la  $\beta$ -glucuronidase (Figure 11)

### SAT1, SAT3, SAT3', SAT2, SAT4 ou toutes SATs

Afin d'obtenir une expression de la SAT dans tous les compartiments subcellulaires (cytosol, mitochondrie ou chloroplaste), on introduira le transgène dans le vecteur approprié décrit dans la figure 12.

Entre les bordures gauche (BG) et droite (BD) du T-DNA est cloné un gène de résistance à la kanamycine (NPTII) codant pour la néomycine phosphotransférase utilisé comme marqueur de sélection pour la transformation du tabac. L'expression du gène NPTII est sous la dépendance du promoteur et du terminateur de la nopaline synthase de *A. tumefaciens*. En aval, le gène de la  $\beta$ -glucuronidase cloné entre les sites uniques *Bam*H1 et *Sac*I, est sous le contrôle du promoteur 35S du virus de la mosaïque du chou-fleur (CaMV) et le signal de polyadénylation du gène de la nopaline synthase du plasmide Ti. L'insertion du gène codant pour une SAT s'effectue entre les sites *Bam*H1 et *Sac*I du vecteur délété du gène de la  $\beta$ -glucuronidase (Figure 12).

### Expression des transgènes dans les graines

Une construction similaire à celle présentée dans les figure 11 ou 12 est réalisée dans le but d'obtenir une expression spécifique du transgène dans les graines. Cette stratégie pourrait être importante puisque les graines composent l'apport principale pour l'alimentation animale. Pour cela le promoteur constitutif de la mosaïque du tabac sera remplacé par un promoteur qui permet une expression spécifique du transgène pendant la mise en place des réserves de la graine.

35

### **Exemple 9. Transformation du tabac**

De jeunes feuilles de plants de tabac (âgés de 3 à 4 semaines) dont la surface est stérilisée avec une solution de javel 10% (V/V) pendant 10 min puis rincée à l'eau

stérile, sont découpées à l'emporte-pièce (30 disques par construction). 20 ml d'une culture de 48 heures d'*Agrobacterium tumefaciens* EHA105 (contenant le vecteur pBI121 modifié selon l'invention) sont centrifugés puis resuspendus dans 4 ml d'une solution de MgSO<sub>4</sub> 10mM. Les disques foliaires sont incubés pendant quelques minutes

5 avec la solution d'agrobactéries puis transférés sur milieu MS (Sigma M-5519) supplémenté avec 0,05 mg/l d'acide  $\alpha$ -naphtalène acétique (NAA, Sigma), 2 mg/l de 6-benzylaminopurine (BAP) et 7 mg/l de phytoagar, pendant 2 à 3 jours. Les disques foliaires sont ensuite transférés sur un milieu identique auquel sont ajoutés 350 mg/l de cefotaxime (bacteriostatique) et 75 mg/l de kanamycine (agent de sélection). Au bout de

10 2 semaines, les disques sur lesquels se sont développés des cals ainsi que de jeunes pousses sont repiqués sur un milieu identique afin d'accélérer la croissance des pousses. Une semaine plus tard, les pousses vertes sont excisées et transférées sur le même milieu sans hormone, afin de permettre le développement des racines, ceci pendant 2 semaines environ, au bout desquelles les jeunes plantes sont transférées en terre et

15 cultivées en serre.

**Exemple 10. Analyse des résultats pour les plantes transgéniques SAT3 et SAT1' (L78443) (forme tronquée de la SAT1 U22964) et témoins.**

L'impact de l'expression de SAT3, SAT1', OTP-SAT3 dans les feuilles ou dans

20 les graines de plants de tabac est analysé au niveau du contenu en composés soufrés, la cystéine et la méthionine (et dérivés comme le S-méthylméthionine ou SMM) et le glutathion. La cystéine et le glutathion sont mis en évidence par la Méthode de Fahey ([33] Fahey, R.C. and Newton, G.L., Methods Enzymol. (1987) 143, 85-96), après dérivatization des composés par le thiolyte (mBBR de Calbiochem) et séparation par

25 chromatographie liquide haute performance (CLHP) [33]. Le dosage de la méthionine libre et du SMM est effectué par les méthodes de dosage des acides aminés libres après extraction et dérivation par l'orthophtaldéhyde et séparation par CLHP ([34] Brunet, P. & al., J. Chrom. (1988) 455, 173-182). La mesure de l'activité de la serine acétyltransférase s'effectue comme décrit dans la méthodologie de dosage de l'O-

30 acétylserine formée, par la technique HPLC, ou par la méthode de couplage en présence d'un excès d'O-acétylserine (thiol) lyase [12], [14]. L'activité du transgène SAT dans les plants transformés (c.a.d *in vivo*) est révélée par le dosage de l'O-acétylsérine produite lors de l'activité de l'enzyme et accumulée transitoirement dans la cellule.

Le dosage de l'O-acétylserine dans les extraits de plantes suit le protocole

35 suivant.

Après broyage des feuilles de tabacs en une fine poudre dans l'azote liquide, les extraits sont repris dans l'acide chlorhydrique 0,1 M (1 ml /100 mg de poudre). Après une période d'incubation de l'ordre de 10 min, les débris sont éliminés par une centrifugation de 15 min à 15.000g. Une fraction du surnageant obtenu, contenant les

acides aminés libres, est dérivatisée pendant 1 min à 25°C en présence d'une solution d'orthophtalaldéhyde (solution de 54 mg d'orthophtalaldéhyde, 10% méthanol, 90% de borate de sodium 400 mM, pH 9.5, et 0.2 ml de  $\beta$ -mercaptoethanol). Les dérivés OPA-acides aminés sont alors séparés par chromatographie en phase inverse sur une colonne  
 5 UPHDO-15M (0.46 x 150 mm, Interchim) connectée à un système HPLC (Waters). Les tampons utilisés pour réaliser l'élution sont, Tampon A : acétate de sodium, 85 mM, pH 4.5 additionné d'acétonitrile 6% final ; tampon B : 60% acétonitrile dans l'eau. La séparation des dérivés s'effectue selon le gradient (1 ml/min): 0 min, 30% B dans A ; 8 min, 60 % B dans A ; 9 min, 80 % B dans A ; 10 min, 100 % B ; 12 min, 100 % B. En  
 10 sortie de colonne, la fluorescence émise par les dérivés est mesurée à 455 nm après excitation à 340 nm (fluorimètre SFM25, Kontron).

Le temps de rétention de l'*O*-acétylsérine dans nos conditions expérimentales est de 9,5 min. L'identité du pic correspondant à l'*O*-acétylsérine est confirmée par une co-  
 élution avec une quantité connue du produit pur. De plus, un deuxième contrôle est  
 15 effectué pour confirmer la position de l'*O*-acétylsérine dans les différentes analyses. Les échantillons, avant l'incubation avec l'OPA, sont préalablement traités par du NaOH 0.2 M finale. Dans ces conditions, la fonction acétate en position OH de la serine est transférée sur la fonction amine et permettant ainsi la formation de la *N*-acétylsérine. Ce dernier composé n'est plus détecté dans nos conditions expérimentales et conduit donc à  
 20 la disparition du pic correspondant initialement à l'*O*-acétylsérine.

Les plants transformés avec le transgène SAT ont été préalablement sélectionnés sur kanamycine et menés à graines. Les plants contrôles (PBI, trois lignées indépendantes contenant le vecteur de transformation et une cassette GUS) sont traités de façon identique. Les analyses des plantes comprennent : 1 ; mise en évidence de  
 25 l'insertion du transgène au niveau du génome par PCR en utilisant les primers 5' et 3' correspondant aux SAT utilisées pour la transformation ; 2, mise en évidence du messenger par une analyse des messagers à l'aide de sondes correspondant aux transgènes SATs utilisés pour la transformation des plantes selon les techniques connues ; 3, mise en évidence de l'activité enzymatique associée à la protéine SAT  
 30 selon les méthodes décrites dans la littérature ([14]) et localisation du transgène ; 4, dosage du produit de la réaction SAT, soit l'*O*-acétylserine (OAS) dans les plantes transformées ; 5, dosage de la cystéine et de ses dérivés directs, le glutathion et la méthionine (et ses dérivés méthylés) ; 5, analyse de la composition en acides aminés totaux des plantes et graines associées à chacun des transgènes obtenus (acides aminés  
 35 libres et liés aux protéines) selon les techniques traditionnelles ; 6 ; analyse de l'impact de la sur-expression de l'activité SAT dans la cellule végétale sur le contenu en activité enzymatique associée à la séquence d'assimilation du soufre (transporteurs de sulfate, ATP-sulfurylase, APS reductase, sulfite reductase et en particulier l'*O*-acétylserine

(thiol) lyase, l'enzyme directement associée à l'activité SAT pour la synthèse de la cystéine ([14]). De plus, les enzymes associées à la séquence de synthèse de la méthionine et du glutathion sont analysées afin de rendre compte de l'impact du contenu en cystéine sur le métabolisme associé à la synthèse du glutathion et de la méthionine.

5 L'expression du gène de la serine acétyltransférase d'*Arabidopsis thaliana* dans le tabac conduit à une augmentation du taux en cystéine, du taux en glutathion et du taux en méthionine dans les tissus des plantes transformées par rapport aux plantes contrôles. En général cette augmentation du contenu en composés soufrés libres est associée à l'expression du transgène dans la cellule végétale (**Figure 13**). La mesure est  
10 effectuée dans les feuilles de 3 plantes différentes de chaque lignée homozygote. L'activité SAT est mesurée par sa capacité à promouvoir la synthèse de cystéine selon le protocole décrit précédemment ([14]).

L'expression du transgène sous le contrôle du promoteur constitutif CaMV conduit à augmenter la capacité (l'activité enzymatique potentielle maximale mesurée *in*  
15 *vitro*) de la SAT d'un facteur 2 à 8 par rapport au niveau mesuré dans les plantes contrôles (plantes transformées avec un vecteur vide). Pour rendre compte de l'activité réelle du transgène SAT, une mesure du contenu en O-acétylserine (OAS libre) a été effectuée. Ainsi, le taux d'OAS dans la cellule végétale (taux moyen de 4 nmoles/ g matière fraîche pour les plantes contrôles, 6 mesures indépendantes) a pu être multiplié  
20 d'un facteur 2 à 10 fois dans les plantes transformées (2 mesures indépendantes). Ainsi, associée à l'augmentation nette de la capacité de l'activité enzymatique de la SAT est associée une augmentation de l' OAS libre intracellulaire qui résulte de l'activité du transgène *in vivo*, et à une augmentation du contenu en cystéine libre dans la plupart des transgènes SAT par rapport aux plants contrôles (**Figure 14**). La teneur en cystéine dans  
25 les plantes contrôle (PBI) et dans les plants T2 de tabac transformés avec une SAT (lignées SAT1' et SAT3) est déterminées comme dérivés de monobromobimane par HPLC pour 3 plantes par lignées ([33]). La teneur en cystéine des lignées transgéniques est augmentée de 2 à 10 fois par comparaison aux plantes contrôle (PBI).

Le contenu en cystéine libre dans la plupart des plantes transgéniques exprimant  
30 une SAT est significativement supérieur de 2 à 10 fois au taux naturel mesuré dans les plantes contrôles PBI (d'une valeur de 5 nmoles /g matière fraîche, moyenne calculée sur trois lignées indépendantes et contenant chacune 5 plantes). Cet impact de l'expression de la SAT est observé dès la génération T1. Par contre aucune corrélation n'a pu être mise en évidence entre le contenu en cystéine (et par ailleurs en OAS libre)  
35 et l'activité SAT des transgènes mesurées *in vitro*. Par contre une corrélation positive et significative a pu être mesurée entre le contenu en OAS cellulaire et le taux de cystéine de la cellule (**Figure 15**). Par comparaison aux plantes contrôle, une augmentation de 3 à 10 fois du niveau d'O-acétylserine libre *in vivo* liée à l'activité du transgène, résulte en une augmentation de 3 à 8 fois du niveau de cystéine dans les plantes. L'analyse a été

réalisée sur des feuilles complètement développées (environ 2 mois) de plantes homozygotes pour le transgène. Les plantes contrôle sont des plantes transformées avec des construits vides (PBI). Une augmentation du contenu en OAS libre cellulaire liée à l'activité du transgène SAT dans les plants transformés est positivement corrélée avec  
 5 une augmentation du contenu en cystéine. Ainsi, une augmentation en moyenne de 6 fois du taux d'OAS libre est associée avec une augmentation de 6-fois le taux en cystéine. La pente associée avec la distribution des points est de  $1,06 \pm 0,09$  (coefficient de régression de 0.67). Elle indique que pour chaque molécule d'OAS accumulée une mole de cystéine est synthétisée. La valeur de cette pente et l'absence de  
 10 plateau observées dans nos conditions expérimentales indiquent que la formation de sulfide (assimilation du sulfate et réduction en sulfide) n'est pas une séquence limitante et que l'activité SAT paraît être le facteur limitant dans la cellule pour la formation de la cystéine (**Figure 1**)

La localisation sub-cellulaire des transgènes SAT1' (forme tronquée de la SAT1) et SAT3 dans les plants de tabacs transformés a pu être précisée par une préparation de  
 15 la fraction chloroplastique des plants transformés présentant l'activité enzymatique la plus élevée par rapport au plants PBI (contrôles). L'activité associée au compartiment chloroplastique est comparée à celle mesurée dans l'extrait total (**Figure 16**). Les valeurs d'activité serine acetyltransférase correspondent à 3 lignées pour les PBI (5  
 20 plantes par lignées), 5 lignées pour la SAT1' et la SAT3 représentées par 5 plantes chacunes. Les colonnes en gris correspondent aux activités mesurées dans l'extrait total réalisé à partir de chacune des lignées, et les colonnes en noir représentent la moyenne des activités mesurées dans chacune des préparations chloroplastiques.

Ces résultats établissent définitivement que la SAT3 est une isoforme de la  
 25 serine acetyltransférase localisée dans le cytosol de la cellule végétale, et que la forme tronquée de la SAT1 (absence de peptide de transit) est aussi localisée dans le compartiment cytosolique. En ce qui concerne la SAT3, ces résultats confirment nos interprétations dérivées de l'analyse de la séquence protéique [12].

Une conséquence directe de l'augmentation du taux de la cystéine cellulaire  
 30 résulte en une synthèse accrue du glutathion et de la méthionine (voir **Figure 1**). Le devenir de la cystéine est multiple et mis à part son incorporation au niveau des protéines, sa participation à la synthèse de composés multiples comme les vitamines (biotine, thiamine, ...et autre dérivés soufrés de la cellule), la cystéine participe aussi à la synthèse du glutathion (tripeptide associé à de nombreux mécanismes de défense de la  
 35 plante et considéré comme réservoir de cystéine) et de la méthionine. En effet, dans les plants transformés avec le transgène SAT le taux de glutathion est directement corrélé à celui de la cystéine et se traduit par une augmentation de 2 à 7 fois le taux naturel mesuré dans les plants contrôles (PBI) (**Figure 17**). Le coefficient de corrélation calculé pour la distribution des points est de 0.92. Une augmentation de 4 fois de la teneur en

cystéine dans les plants de tabac transgénique surexprimant la SAT résulte en une augmentation de 3 à 4 fois du niveau de glutathion. L'analyse a été effectuée avec des feuilles pleinement développées (environ 2 mois) de plantes homozygotes pour le transgène. Les plantes contrôles sont des plantes transformées avec des construits vides.

5 Ce résultat indique que la cystéine est le facteur limitant de la synthèse du glutathion dans la cellule végétale. Donc, indirectement toute modification du taux de serine acétyltransférase dans la cellule aura pour conséquence, via l'augmentation du taux de cystéine, une augmentation du contenu en glutathion intra-cellulaire. Ce résultat implique que les plantes transgéniques obtenues ont acquis des propriétés de résistances  
10 aux stress par rapport aux plantes contrôles (PBI). Cet aspect a été observé récemment ([34], Blaszczyk A. & al., 1999, The Plant Journal 20, 237-243). De plus, le contenu en cystéine et glutathion considéré comme un réservoir entraîne une disponibilité accrue lors de la synthèse de polypeptides riches en cystéines (par exemples pour la résistances aux attaques phytopathogènes), et riches cystéine et en méthionine (pour l'alimentation  
15 animales)

L'augmentation de la cystéine dans la cellule végétale entraîne aussi une augmentation du contenu relatif en méthionine (**Figure 18**). Par contre, à l'opposé du résultats observés pour le glutathion, la courbe présente un plateau qui semble indiquer l'existence d'un autre site de contrôle qui limiterait la synthèse de la méthionine. De  
20 plus, l'homocystéine, issue de la voie de transsulfuration, et le précurseur soufrés à la synthèse de la cystéine ne semble pas s'accumuler. Cette observation indique donc que le pool de folates de la cellule végétale, indispensable pour la méthylation et la formation de la méthionine n'est pas un facteur limitant. Cette limitation se situerait donc en aval de la cystéine et en amont de l'homocystéine. Elle concerne la synthèse du  
25 précurseur carboné pour la synthèse de la méthionine qui dérive de l'aspartate (*O*-phosphohomoserine et/ou cystathionine). Le niveau de l'aspartokinase (la première enzyme de la voie de l'aspartate pour la synthèse de la Lysine, thréonine et méthionine) est contrôlé par plusieurs effecteurs comme la thréonine, et le *S*-adenosylmethionine (SAM) issue de la synthèse de la méthionine ([3]). La cystathionine  $\gamma$ -synthase (voir  
30 **Figure 1**) est directement régulé au niveau transcriptionnelle [3] et plus exactement la méthionine ou l'un de ses dérivés contrôle la stabilité de son messenger [4]). Le plateau maximale obtenu dans nos conditions expérimentales est de l'ordre de  $39 \pm 7$  nmoles /g matière fraîche de méthionine ce qui correspond à une multiplication du taux naturel moyen de l'ordre de  $6 \pm 2$  nmole per g matière fraîche (contrôle PBI). La valeur  
35 maximale obtenue pour la méthionine nécessite une augmentation du contenu en cystéine de la cellule de 4 à 5 fois son taux maximale. Le coefficient de régression est de 0.50.

De plus, l'augmentation de la méthionine dans les cellules conduit à multiplier le taux de *S*-methyilméthionine (SMM) de 2 à 10 fois selon les plantes. Le SMM dérive

directement de la méthylthion de la méthionine en présence de *S*-adenosylméthionine. Ce composé est important pour la cellule et est une forme de transport des groupements méthyls (de méthionine) dans la plante. En effet, en présence d'une molécule d'homocystéine (le précurseur soufré à la synthèse de la méthionine et dérivant de la cystéine), le SMM permet la synthèse de deux molécules de méthionine ([3], [35], Bourgis & al., 1999, *Plant Cell* 11, 1485-1497). Il pourrait donc avoir un rôle primordial lors de la synthèse des protéines de réserves dans la graine. De plus, le SMM est le précurseur direct pour la synthèse de composé comme le 3-diméthylsulfoniopropionate impliqué dans la résistance des plantes aux stress salins ([36], Hanson A.D. & al., 1994, *Plant Physiol.* 105, 103-110). Une telle approche présente de multiples conséquences en particulier pour augmenter les potentialités des plantes sur des sols riches en sels.

#### Evidence pour un rôle régulateur de la voie d'assimilation du sulfate *in vivo*.

La serine acetyltransferase est considéré comme un facteur limitant pour l'assimilation du soufre et la synthèse de la cystéine. Son rôle chez les bactéries est important puisque le produit de la réaction, (*O*-acetylserine, OAS) ou son dérivé (le *N*-acetylserine) est l'effecteur qui module l'expression des gènes de la séquence d'assimilation du soufre comme: 1, le transport du sulfate, 2, l'ATP sulfurylase, 3, l'APS kinase, et 4, la PAPS reductase ([37], Kredich N.M., 1987, in *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology, pp. 419-428). Chez les plantes, un rôle de l'OAS dans la modulation de l'expression de plusieurs gènes a pu être mise en évidence et concerne les transporteurs de sulfates, ([38], Smith F.W. & al. 1997, *The Plant Journal* 12, 875-884; [39], Hawkesford M.J. & al. 1995, *Z. Pflanzenernähr. Bodenk.* 158, 55-57; [40], Clarkson D.T. & al. 1999, *Plant Physiol. Biochem.* 37, 283-290), l'ATP sulfurylase ([39-40]) et l'APS reductase ([41], Neuenschwander U. & al. 1991, *Plant Physiol.*, 97, 253-258). Le rôle de l'activité serine acetyltransferase dans la modulation des gènes a été postulé d'après les caractéristiques cinétiques du complexe cystéine synthase (bi-enzyme complexe composé de la serine acetyltransferase et de l'*O*-acetylserine (thiol) lyase) ([41], Droux & al. In *Sulphur and Nutrition in Plants*, sous Presse), et a conduit à décrire un modèle pour rendre compte du mécanisme de régulation des gènes. Le rôle de l'OAS est aussi déterminant dans la régulation de l'expression des gènes lors de la formation des graines ([42], Kim H. & al., 1999, *Planta* 209, 282-289).

Dans les plantes transgéniques qui sur-expriment une SAT dans le cytosol, une augmentation transitoire en OAS a pu être mise en évidence (augmentation de 2 à 10 fois son taux naturel, voir **figure 15**). Parallèlement, dans la plupart des plantes transgéniques, une augmentation de l'activité OASTL a été mesurée (**Figure 19**). Cette augmentation de 2 à 5 fois par rapport à l'activité mesurée dans les contrôles PBI, ne concerne que l'activité associée au chloroplaste. De plus, dans un western Blot, le signal

observée est plus important dans la plupart des lignées transgéniques (**Figure 20**) indiquant que l'augmentation de l'activité correspond à une induction de la synthèse *de novo* de l'OASTL. Ce résultat original correspond à la première démonstration du rôle de l'OAS (*in planta*) dans la modulation des gènes de la séquence d'assimilation du sulfate, en particulier pour l'OASTL chloroplastique.

Pour la **figure 20**, la quantité équivalente de protéine (0.150 mg) est soumise à un SDS-PAGE (12%) et après séparation des protéines, celles-ci sont transférées sur une membrane de nitrocellulose. La présence de l'OASTL est révélée par une incubation avec des anticorps dirigés contre l'OASTL chloroplastique de feuille d'épinards ([7]).

La sur-expression de la SAT dans la cellule végétale conduit donc à augmenter la capacité à synthétiser la cystéine dans le chloroplaste. Il est donc permis de supposer que l'expression des gènes codant pour les enzymes de la voie d'assimilation et de réduction du sulfate (transporteur de sulfate, ATP sulfurylase, APS reductase, sulfite reductase) est aussi modulée comme l'OASTL (et références [38-41]).

L'augmentation du contenu intracellulaire en OAS (qui dérive de l'activité de la SAT) signale un état de stress soufrés (absence de soufre réduit suffisant) artificielle à la cellule (dans les plantes transformées) qui conduit à induire les enzymes de la voie d'assimilation du sulfate.

**Impact de l'augmentation en cystéine dans la cellule sur le contenu générale en acides aminés.** Cette augmentation en composés soufrés s'accompagne d'une augmentation du contenu en acides aminés essentiels comme la thréonine, l'isoleucine, et la lysine (leur contenu est multiplié par 2 en moyenne). Par contre le taux de glutamate est divisé par 2 comme celui de l'aspartate. Cette dernière observation est directement liée à l'augmentation du contenu en THR, LYS et ILE. Toutes les augmentations en acides aminés sont corrélées avec une augmentation en l'activité de la serine acétyltransférase (SAT3, ou SAT1') dans le cytosol. De plus, l'augmentation de ces composés soufrés conduit à améliorer le rapport nutritionnelle N/S de la plante (sur la base des acides aminé libres). Il se traduit par une baisse de ce rapport relatif due à l'enrichissement en composé soufrés totaux (cysteine, méthionine, SMM et glutathion). Ce facteur est important car il conditionne le contenu des graines en polypeptides et conduit à un enrichissement (ou à un appauvrissement si le rapport N/S est trop élevé) des protéines de réserves riches en acides aminé soufrés au détriment des polypeptides pauvres en ces composés.

#### **Exemple 11 Analyse des plants transgéniques OTP-SAT3 (OTP-SAT1')**

L'analyse des transformants au stade TO des plantes transgéniques exprimant une SAT insensible à la cystéine (ici comme exemple la SAT3 ou SAT1', forme tronquée de la SAT1 U22964), dans les feuilles ou dans les graines (sous le contrôle d'un

promoteur graine spécifique), révèle une augmentation de la teneur en cystéine libre, mais aussi du glutathion (2.6 fois le taux naturel), et en méthionine. Les plantes exprimant ces mêmes isoformes dans le cytosol sous le contrôle d'un promoteur graine spécifique révèlent un taux de composé soufrés supérieurs aux plantes témoins.

5

**Exemple 12 Analyse des résultats pour les plantes transgéniques SAT1 (CDNA U22964 ou SAT1<sub>ju</sub>, forme avec peptide de transit) et témoins.**

L'impact d'une expression de la serine acetyltransferase dans les mitochondries a été analysé en transformant les plantes avec la construction (**Figure 12**) contenant la  
 10 séquence entière de la SAT1. L'analyse des plantes au stade TO permet de mettre en évidence une augmentation en cystéine libre dans la cellule (**Figure 21**). L'analyse est effectuée sur les feuilles formées avant l'apparition de la hampe florale. Les quatorze lignées présentent une multiplication du taux de cystéine de 2 à 6 fois par comparaison avec le plant contrôle (PBI).

15 L'augmentation en cystéine s'accompagne d'un effet général sur le contenu en composé soufrés avec une multiplication par 4 du contenu en glutathion dans la cellule (**Figure 22**). A l'opposé d'une expression de la SAT dans le compartiment cytosolique, l'allure générale de la distribution des valeurs dans les différentes lignées présente un plateau qui indiquerait une limitation dans la synthèse du glutathion. Cette limitation  
 20 peut concerner le taux de glutamate et/ou de glycine, ou un contrôle du glutathion sur sa propre synthèse (rétro-inhibition de l'une des enzymes participant à la synthèse du glutathion, enzyme E6 et/ou E7 voir **Figure 1**).

De même le contenu en méthionine est multipliée par 2 à 3 fois par rapport au taux naturel mesuré dans les plants contrôles.

### Revendications

1. Procédé pour augmenter la production de cystéine, glutathion, méthionine et leurs dérivés soufrés par les cellules végétales et les plantes, ledit procédé consistant à surexprimer une SAT dans les cellules végétales, et les plantes contenant les dites cellules végétales.
2. Procédé selon la revendication 1, caractérisé en ce que la SAT surexprimée dans les cellules végétales est une SAT sensible à la cystéine. 3.
- Procédé selon la revendication 2, caractérisé en ce que la SAT est une SAT de plante ou une SAT native d'origine bactérienne.
4. Procédé selon la revendication 1, caractérisé en ce que la SAT surexprimée dans les cellules végétales est une SAT insensible à la cystéine.
5. Procédé selon la revendication 4, caractérisé en ce que la SAT est une SAT de plante ou une SAT d'origine bactérienne ou de plante mutée, rendue insensible à la cystéine par mutagenèse.
6. Procédé selon l'une des revendications 1 à 5, caractérisé en ce que la SAT est surexprimée dans le cytoplasme des cellules végétales.
7. Procédé selon la revendication 6, caractérisé en ce que la SAT est une SAT d'origine bactérienne.
8. Procédé selon la revendication 6, caractérisé en ce que la SAT est une SAT cytoplasmique de plante, en particulier d'*Arabidopsis thaliana*.
9. Procédé selon la revendication 8, caractérisé en ce que la SAT est la SAT 3, représentée le SEQ ID NO 1.
10. Procédé selon la revendication 6, caractérisé en ce que la SAT est une SAT de plante non cytoplasmique amputée de son ou ses signaux d'adressage vers des compartiments cellulaires différents du cytoplasme.
11. Procédé selon la revendication 10, caractérisé en ce que la SAT est la SAT 1' représentée par la SEQ ID NO 2.
12. Procédé selon l'une des revendications 1 à 5, caractérisé en ce que la SAT est surexprimée dans les mitochondries.
13. Procédé selon la revendication 12, caractérisé en ce que la SAT est surexprimée dans le cytoplasme sous la forme d'une protéine de fusion peptide signal/SAT, la SAT mature fonctionnelle étant libérée à l'intérieur des mitochondries.
14. Procédé selon la revendication 13, caractérisé en ce que le peptide signal d'adressage mitochondrial est constitué par au moins un peptide signal d'une protéine végétale naturelle à localisation mitochondriale, comme par exemple le peptide signal de la SAT1 représenté par les acides aminés 1 à 63 sur la SEQ ID NO 3..
15. Procédé selon la revendication 13, caractérisé en ce que la SAT est une SAT mitochondriale d'origine végétale, en particulier d'*Arabidopsis thaliana*.

16. Procédé selon la revendication 15, caractérisé en ce que la SAT est la SAT1 représentée par la SEQ ID NO 3.
17. Procédé selon la revendication 6, caractérisé en ce que la SAT est surexprimée dans les chloroplastes des cellules végétales.
- 5 18. Procédé selon la revendication 17, caractérisé en ce que la SAT est surexprimée dans les chloroplastes par intégration dans l'ADN chloroplastique des cellules végétales d'un gène chimère comprenant une séquence d'ADN codant pour ladite SAT, sous le contrôle d'éléments de régulations en 5' et 3' fonctionnels dans les chloroplastes.
- 10 19. Procédé selon la revendication 17, caractérisé en ce que la SAT est surexprimée dans le cytoplasme sous la forme d'une protéine de fusion peptide de transit/SAT, la SAT mature fonctionnelle étant libérée à l'intérieur des chloroplastes.
20. Procédé selon la revendication 19, caractérisé en ce que la SAT est homologue du peptide de transit.
- 15 21. Procédé selon la revendication 20, caractérisé en ce que la SAT est une SAT chloroplastique d'origine végétale, en particulier d'*Arabidopsis thaliana*.
22. Procédé selon la revendication 21, caractérisé en ce que la SAT est la SAT2 ou la SAT4 représentée par la SEQ ID NO 5 ou 6 respectivement.
23. Procédé selon la revendication 19, caractérisé en ce que la SAT est
- 20 hétérologue du peptide de transit.
24. Procédé selon la revendication 13, caractérisé en ce que la SAT est une SAT cytoplasmique d'origine végétale ou une SAT d'origine bactérienne, telle que définie dans l'une des revendications 3 à 5 ou 9 à 11.
25. Procédé selon l'une des revendications 23 ou 24, caractérisé en ce que le
- 25 peptide de transit est un peptide de transit d'une autre protéine à localisation plastidiale.
26. Procédé selon la revendication 25, caractérisé en ce que le peptide de transit est constitué par le peptide de transit d'une EPSPS de plante ou le peptide de transit d'une ssu RuBisCO de plante.
27. Procédé selon l'une des revendications 25 ou 26, caractérisé en ce que le
- 30 peptide de transit comprend un peptide de transit d'une protéine végétale à localisation plastidiale et une partie de séquence de la partie mature N-terminale d'une protéine à localisation plastidiale entre la partie C-terminale du peptide de transit et la partie N-terminale de la SAT.
28. Procédé selon la revendication 27, caractérisé en ce que la partie de
- 35 séquence comprend généralement moins de 40 acides aminés de la partie N-terminale de la protéine mature, de préférence moins de 30 acides aminés, plus préférentiellement entre 15 et 25 acides aminés.
29. Procédé selon l'une des revendications 27 ou 28, caractérisé en ce que le peptide de transit comprend entre la partie C-terminale de la partie N-terminale de la

protéine mature et la partie N-terminale de la SAT un deuxième peptide de transit d'une protéine végétale à localisation plastidiale.

30. Procédé selon la revendication 29, caractérisé en ce que le peptide de transit est un peptide de transit optimisé (OTP) constitué par la fusion d'un premier peptide de transit, avec une partie de séquence de la partie mature N-terminale d'une protéine à localisation plastidiale, laquelle est fusionnée avec un deuxième peptide de transit.

31. Protéine de fusion peptide de transit/SAT, caractérisé en ce que la SAT est hétérologue du peptide de transit.

32. Protéine de fusion selon la revendication 31, telle que définie dans les revendications 24 à 30.

33. Séquence d'acide nucléique codant pour une protéine de fusion peptide de transit/SAT selon l'une des revendications 31 ou 32.

34. Gène chimère comprenant une séquence codante ainsi que des éléments de régulation en position 5' et 3' hétérologues pouvant fonctionner dans un organisme hôte, caractérisé en ce que la séquence codante comprend au moins une séquence d'acide nucléique codant pour une SAT.

35. Gène chimère selon la revendication 34, caractérisé en ce que l'organisme hôte est choisi parmi les bactéries, par exemple *E. coli*, les levures, en particulier des genres *Saccharomyces* ou *Kluyveromyces*, *Pichia*, les champignons, en particulier *Aspergillus*, les baculovirus, ou les cellules végétales et les plantes.

36. Gène chimère selon la revendication 35, caractérisé en ce que l'organisme hôte est une cellule végétale ou une plante la contenant.

37. Gène chimère selon la revendication 36, caractérisé en ce que l'élément de régulation en 5' comprend les séquences de régulation promotrice dans cellules végétales et les plantes, choisi parmi les promoteurs s'exprimant dans les feuilles des plantes, les promoteurs constitutifs, ou les promoteurs lumière dépendants d'origine bactérienne, virale ou végétale

38. Gène chimère selon la revendication 36, caractérisé en ce que l'élément de régulation en 5' comprend les séquences de régulation promotrice dans les cellules végétales et les plantes, choisi parmi les promoteurs spécifiques des graines.

39. Gène chimère selon la revendication 38, caractérisé en ce que le promoteur est choisi parmi les promoteurs de la napine, de la phaseoline, de la glutenine, de la zéine, de l'héliantinine, de l'albumine ou de l'oléosine.

40. Gène chimère selon l'une des revendications 34 à 39, caractérisé en ce que la séquence d'acide nucléique codant pour une SAT code pour une SAT telle que définie dans les revendications 2 à 30.

41. Gène chimère selon l'une des revendications 34 à 39, caractérisé en ce que la séquence d'acide nucléique codant pour une SAT est la séquence d'acide

nucléique selon la revendication 33.

42. Vecteur de clonage et/ou d'expression pour la transformation d'un organisme hôte caractérisé en ce qu'il contient au moins un gène chimère tel que défini selon l'une des revendications 34 à 41.

5 43. Procédé de transformation des organismes hôtes caractérisé en ce que l'on intègre dans le génome dudit organisme hôte au moins une séquence d'acide nucléique selon la revendication 33 ou un gène chimère selon l'une des revendications 34 à 41.

10 44. Procédé selon la revendication 43, au moyen du vecteur selon la revendication 42.

45. Procédé selon l'une des revendications 43 ou 44, caractérisé en ce que l'organisme hôte est choisi parmi les bactéries, par exemple *E. coli*, les levures, en particulier des genres *Saccharomyces* ou *Kluyveromyces*, *Pichia*, les champignons, en particulier *Aspergillus*, les baculovirus, ou les cellules végétales et les plantes.

15 46. Procédé selon la revendication 45, caractérisé en ce que l'organisme hôte est une cellule végétale ou une plante la contenant.

47. Procédé selon la revendication 46, caractérisé en ce que la plante est régénérée à partir d'une cellule végétale transformée.

20 48. Procédé selon la revendication 47, caractérisé en ce que l'organisme hôte est une plante monocotylédone, en particulier choisie parmi les céréales, la canne à sucre, le riz et le maïs, ou une plante dicotylédone, en particulier choisie parmi le tabac, la soja, le colza, le coton, la betterave et le trèfle.

25 49. Organisme hôte transformé, caractérisé en ce qu'il comprend au moins une séquence d'acide nucléique selon la revendication 33 ou un gène chimère selon l'une des revendications 34 à 41.

50. Organisme hôte selon la revendication 49, caractérisé en ce qu'il est obtenu par le procédé selon l'une des revendications 43 à 48.

30 51. Cellule végétale, caractérisée en ce qu'elle comprend au moins une séquence d'acide nucléique selon la revendication 33 ou un gène chimère selon l'une des revendications 34 à 41.

52. Plante génétiquement modifiée, caractérisée en ce qu'elle comprend au moins une cellule végétale selon la revendication 51.

53. Plante selon la revendication 52, caractérisé en ce que la plante est régénérée à partir d'une cellule végétale selon la revendication 51.

35 54. Plante génétiquement modifiée, caractérisée en ce qu'elle est issue de la culture et/ou du croisement des plantes régénérées selon la revendication 53.

55. Plante génétiquement modifiée selon l'une des revendications 52 à 54, caractérisée en ce qu'elle est une plante monocotylédone, en particulier choisie parmi

les céréales, la canne à sucre, le riz et le maïs, ou une plante dicotylédone, en particulier choisie parmi le tabac, la soja, le colza, le coton, la betterave et le trèfle.

56. Plante génétiquement modifiée selon l'une des revendications 52 à 55, caractérisée en ce qu'elle comprend d'autres gènes d'intérêt.

5 57. Plante génétiquement modifiée selon la revendication 56, caractérisée en ce qu'elle comprend au moins un autre gène modifiant la teneur et la qualité des protéines de ladite plante, en particuliers dans les feuilles et/ou les graines.

58. Plante génétiquement modifiée selon l'une des revendications 56 ou 57, caractérisée en ce que le gène code pour une protéine enrichie en acides aminés soufrés.

10 59. Graines des plantes génétiquement modifiées selon l'une des revendications 52 à 58.

**Procédé pour augmenter la teneur en composés soufrés et notamment en cystéine,  
méthionine et glutathion chez les plantes et plantes obtenues**

**RHONE-POULENC AGRO**

**Abrégé Descriptif**

La présente invention concerne un procédé pour augmenter la production de cystéine, de méthionine, de glutathion et de leurs dérivés par les cellules végétales et les plantes, ledit procédé consistant à surexprimer une SAT dans les cellules végétales, et les plantes contenant les dites cellules végétales.

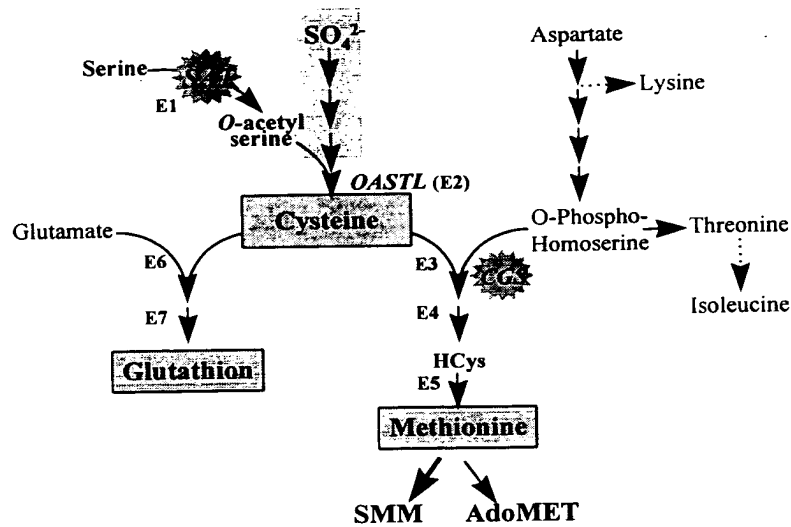


Figure 1 : Séquence illustrant la voie de synthèse de la cystéine et des dérivés soufrés (glutathion et méthionine).

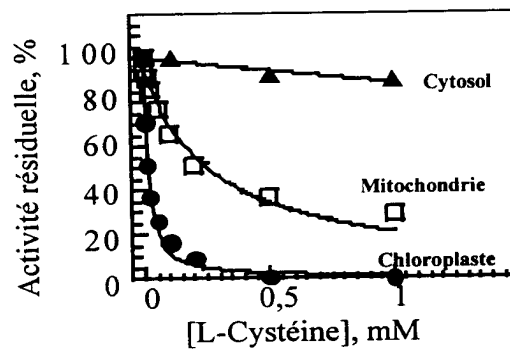


Figure 2 : Effet de la cystéine sur les activités sérine acétyltransférase de pois (*Pisum sativum*).

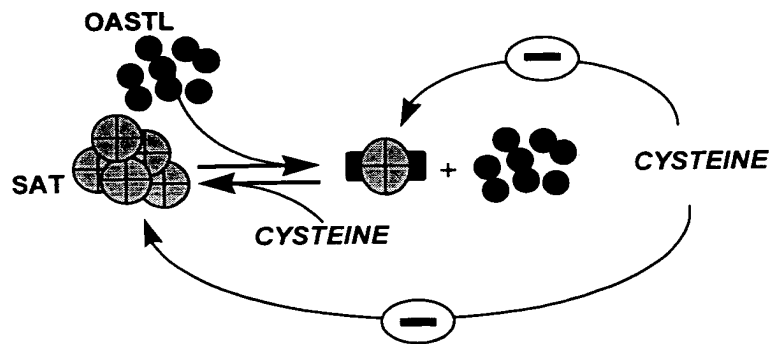


Figure 3 : Modèle de l'inhibition de la sérine acétyltransférase chloroplastique.

ATG	GCA	ACA	TGC	ATA	GAC	ACA	TGC	CGA	ACC	GGT	AAT	ACC	CAA	GAC	GAT	48
GAT	TCC	CGG	TTC	TGT	TGC	ATC	AAG	AAT	TTC	TTT	CGA	CCC	GGT	TTC	TCT	96
GTA	AAC	CGG	AAG	ATT	CAC	CAC	ACC	CAA	ATC	GAA	GAT	GAC	GAT	GAT	GTC	144
TGG	ATC	AAG	ATG	CTT	GAA	GAA	GCC	AAA	TCC	GAT	GTT	AAA	CAA	GAA	CCC	192
ATT	TTA	TCA	AAC	TAC	TAC	TAC	GCT	TCG	ATC	ACA	TCT	CAT	CGA	TCT	TTA	240
GAG	TCT	GCT	TTA	GCT	CAC	ATC	CTC	TCC	GTA	AAG	CTC	AGC	AAT	TTA	AAC	288
CTA	CCA	AGC	AAC	ACA	CTC	TTC	GAA	CTG	TTC	ATA	AGC	GTT	TTA	GAA	GAA	336
AGC	CCT	GAG	ATC	ATC	GAA	TCC	ACG	AAG	CAA	GAT	CTT	ATA	GCA	GTC	AAA	384
GAA	AGA	GAC	CCA	GCT	TGT	ATA	AGC	TAC	GTT	CAT	TGC	TTC	TTG	GGC	TTC	432
AAA	GGC	TTC	CTC	GCT	TGT	CAA	GCT	CAT	CGA	ATA	GCT	CAT	ACC	CTC	TGG	480
AAA	CAG	AAC	AGA	AAA	ATC	GTA	GCT	TTA	TTG	ATC	CAA	AAC	AGA	GTA	TCA	528
GAA	TCT	TTC	GCC	GTC	GAT	ATT	CAT	CCC	GGA	GCG	AAG	ATC	GGA	AAA	GGG	576
ATT	CTT	TTA	GAC	CAT	GCG	ACG	GGC	GTG	GTG	ATC	GGA	GAG	ACG	GCG	GTG	624
GTT	GGA	GAC	AAT	GTT	TCG	ATT	CTA	CAC	GGA	GTG	ACC	TTG	GGA	GGA	ACA	672
GGG	AAA	CAG	AGT	GGT	GAT	CGG	CAT	CCG	AAG	ATT	GGT	GAT	GGT	GTG	TTG	720
ATT	GGA	GCT	GGG	AGT	TGT	ATA	TTG	GGG	AAT	ATA	ACA	ATC	GGT	GAG	GGA	768
GCT	AAG	ATT	GGA	TCA	GGG	TCG	GTG	GTG	GTT	AAG	GAT	GTG	CCG	GCG	CGT	816
ACG	ACG	GCG	GTT	GGA	AAT	CCG	GCG	AGG	TTG	ATT	GGT	GGG	AAA	GAG	AAT	864
CCG	AGA	AAA	CAT	GAT	AAG	ATT	CCT	TGT	CTG	ACT	ATG	GAC	CAG	ACA	TCG	912
TAT	TTA	ACC	GAG	TGG	TCT	GAT	TAT	GTG	ATT	TAA						945

**Figure 4: Séquence nucléotidique et peptidique du gène de l'isoforme SAT 3 (L34076) d'*A. thaliana***

		M	P	P	A	G	E	L	R	H	Q	S	P	S	K	14
		ATG	CCA	CCG	GCC	GGA	GAA	CTC	CGA	CAT	CAA	TCT	CCA	TCA	AAG	42
1	N	L	S	S	V	T	Q	S	D	E	A	E	A	A	S	30
2	GAG	AAA	CTA	TCT	TCC	GTT	ACC	CAA	TCC	GAT	GAA	GCA	GAA	GCA	GCG	90
3	A	I	S	A	A	A	A	D	A	E	A	A	S	L	N	46
4	GCA	GCG	ATA	TCT	GCG	GCA	GCT	GCA	GAT	GCG	GAA	GCT	GCC	GGA	TTA	138
5	T	Q	I	K	A	E	A	R	R	D	A	E	A	E	P	62
6	ACA	CAG	ATC	AAG	GCG	GAA	GCT	CGC	CGT	GAT	GCT	GAG	GCG	GAG	CCA	186
7	A	S	Y	L	Y	S	T	I	L	S	H	S	S	L	E	78
8	TTA	GCT	AGC	TAT	CTA	TAT	TCG	ACG	ATT	CTT	TCT	CAT	TCG	TCT	CTT	234
9	S	I	S	F	H	L	G	N	K	L	O	S	S	T	L	94
10	CGA	TCT	ATC	TCG	TTT	CAT	CTA	GGA	AAC	AAG	CTT	TGT	TCC	TCA	ACG	282
11	S	T	L	L	Y	D	L	F	L	H	T	F	S	S	D	110
12	TTA	TCC	ACA	CTT	TTA	TAC	GAT	CTG	TTC	TTA	AAC	ACT	TTT	TCC	TCC	330
13	S	L	R	N	A	T	V	A	D	L	R	A	A	R	V	126
14	CCT	TCT	CTT	CGT	AAC	GCC	ACC	GTC	GCA	GAT	CTA	CGC	GCT	GCT	CGT	378
15	R	D	P	A	C	I	S	F	S	H	D	L	N	Y	N	142
16	CGT	GAT	CCT	GCT	TGT	ATC	TCG	TTC	TCT	TGT	CTC	CTC	AAT	TAC	AAA	426
17	S	F	L	A	I	Q	R	H	R	V	S	H	K	L	W	158
18	GGC	TTC	TTA	GCT	ATT	CAG	GCG	CAT	CGT	GTA	TCA	CAC	AAG	CTA	TGG	474
19	Q	S	R	K	P	L	A	L	A	L	I	S	R	I	S	174
20	CAA	TCA	CGG	AAG	CCA	TTA	GCA	TTA	GCT	CTA	CAC	TCA	AGA	ATC	TCC	522
21	T	F	A	V	D	L	H	F	A	A	K	I	G	K	G	130
22	GTA	TTC	GCT	GTT	GAT	ATC	CAT	CCA	GCA	GCG	AAG	ATC	GGA	AAA	GGG	570
23	L	D	H	A	T	G	V	V	V	G	E	T	A	/	I	206
24	CTT	CTA	GAC	CAC	GCA	ACC	GGA	GTT	GTA	GTC	GGA	GAA	ACA	GCG	GTG	618
25	N	N	V	S	I	L	H	H	V	T	L	G	G	T	G	222
26	GGG	AAC	AAT	GTT	TCA	ATC	CTT	CAC	CAT	GTG	ACA	CTA	GGT	GGA	ACA	666
27	K	A	C	G	D	R	H	P	K	I	G	D	G	C	L	238
28	AAA	GCT	TGT	GGA	GAT	AGA	CAT	CCG	AAG	ATC	GGT	GAC	GGT	TGT	TTG	714
29	G	A	G	A	T	I	L	G	N	V	K	I	G	A	G	254
30	GGA	GCT	GGA	GCG	ACT	ATT	CTT	GGA	AAT	GTG	AAG	ATT	GGT	GCA	GGT	762
31	K	V	G	A	G	S	V	V	L	I	D	V	P	C	R	270
32	AAA	GTA	GGA	GCT	GGT	TCT	GTT	GTG	CTG	ATT	GAC	GTG	CCT	TGT	CGA	810
33	T	A	V	G	N	P	A	R	L	V	G	G	K	E	K	286
34	ACT	GCG	GTT	GGG	AAT	CCG	GCG	AGA	CTT	GTC	GGA	GGG	AAA	GAG	AAG	858
35	L	H	D	E	E	C	P	G	E	S	M	D	H	T	S	302
36	ACG	ATT	CAT	GAT	GAG	GAA	TGT	CCT	GGA	GAA	TCG	ATG	GAT	CAT	ACT	906
37	F	I	S	E	W	S	D	Y	I	I	...					312
38	TTC	ATC	TCG	GAA	TGG	TCA	GAT	TAC	ATC	ATA	TAA					939

**Figure 5: Séquence nucléotidique et peptidique du gène de l'isoforme SAT3' (U30298) d'*A. thaliana***

M	A	A	C	T	D	T	C	R	T	G	R	P	C	T	15
ATG	GCT	GCG	TGC	ATC	GAC	ACC	TGC	CGC	ACT	GGT	AAA	CCC	CAG	ATT	45
S	T	R	D	S	S	K	H	H	D	D	E	S	G	F	30
TCT	CCT	CGC	GAT	TCT	TCT	AAA	CAC	CAC	GAC	GAT	GAA	TCT	GGC	TTT	90
R	V	M	N	Y	F	R	Y	P	D	R	S	S	F	N	45
CGT	TAC	ATG	AAC	TAC	TTC	CGT	TAT	CCT	GAT	CGA	TCT	TCC	TTC	AAT	135
S	T	Q	T	K	T	L	H	T	R	P	L	L	E	D	60
GGA	ACC	CAG	ACC	AAA	ACC	CTC	CAT	ACT	CGT	CCT	TTG	CTT	GAA	GAT	180
T	D	R	D	A	E	V	D	D	T	M	A	K	T	R	75
CTC	GAT	CGC	GAC	GCT	GAA	GTC	GAT	GAT	GTT	TGG	GCC	AAA	ATC	CGA	225
T	R	A	K	S	D	T	A	K	E	P	T	V	S	A	90
GAA	GAG	GCT	AAA	TCT	GAT	ATC	GCC	AAA	GAA	CCT	ATT	GTT	TCC	GCT	270
V	Y	H	A	S	T	V	S	Q	R	S	L	E	A	A	105
TAT	TAT	CAC	GCT	TCG	ATT	GTT	TCT	CAG	CGT	TCG	TTG	GAA	GCT	GCG	315
T	D	N	T	L	S	V	K	L	S	N	L	N	L	P	120
TTG	GCG	AAT	ACT	TTA	TCT	GTT	AAA	CTC	AGC	AAT	TTG	AAT	CTT	CCA	360
T	N	T	L	F	D	L	F	S	G	V	T	Q	G	N	135
AGC	AAC	ACG	CTT	TTC	GAT	TTG	TTC	TCT	GGT	GTT	CTT	CAA	GGA	AAC	405
T	D	T	V	E	S	V	K	T	T	T	L	A	T	K	150
CCA	GAT	ATT	GTT	GAA	TCT	GTC	AAG	CTA	GAT	CTT	TTA	GCT	GTT	AAG	450
F	R	D	P	A	C	T	S	V	T	H	C	F	L	H	165
GAG	AGA	GAT	CCT	GCT	TGT	ATA	AGC	TAC	GTT	CAT	TGT	TTC	CTT	CAC	495
F	K	G	F	L	A	C	Q	A	H	R	T	A	H	E	180
TTT	AAA	GGC	TTC	CTC	GCT	TGT	CAA	GCG	CAT	CGT	ATT	GCT	CAT	GAG	540
T	W	T	Q	D	R	K	T	L	A	L	L	T	Q	N	195
CTT	TGG	ACT	CAG	GAC	AGA	AAA	ATC	CTA	GCT	TTG	TTG	ATC	CAG	AAC	585
R	V	S	E	A	F	A	V	D	F	H	P	G	A	K	210
AGA	GTC	TCT	GAA	GCC	TTC	GCT	GTT	GAT	TTC	CAC	CCT	GGA	GCT	AAA	630
T	C	T	G	I	L	L	D	H	A	T	A	I	V	T	325
ATC	GGT	ACC	GGG	ATT	TTG	CTA	GAC	CAT	GCT	ACG	GCT	ATT	GTG	ATC	675
G	E	T	A	V	V	G	N	N	V	S	I	L	H	N	240
GGT	GAG	ACG	GCG	GTT	GTG	GGG	AAC	AAT	GTT	TCG	ATT	CTC	CAT	AAC	720
V	T	L	G	G	T	G	K	Q	C	G	D	R	H	P	255
GTT	ACG	CTT	GGA	GGA	ACG	GGG	AAA	CAG	TGT	GGA	GAT	AGG	CAC	CCG	765
N	I	G	D	G	V	L	T	G	A	G	T	C	T	L	270
AAG	ATT	GGC	GAT	GGG	GTT	TTG	ATT	GGA	GCT	GGG	ACT	TGT	ATT	TTG	810
G	N	T	T	I	G	E	G	A	K	T	G	A	G	S	285
GGG	AAT	ATC	ACG	ATT	GGT	GAA	GGA	GCT	AAG	ATT	GGT	GCG	GGG	TCG	855
T	W	L	K	D	V	P	P	R	T	T	A	V	G	N	300
GTG	GTG	TTG	AAA	GAC	GTG	CCG	CCG	CGT	ACG	ACG	GCT	GTT	GGA	AAT	900
P	A	R	L	L	G	G	K	D	N	P	K	T	H	D	315
CCG	GCG	AGG	TTG	CTT	GGT	GGT	AAA	GAT	AAT	CCG	AAA	ACG	CAT	GAC	945
K	I	P	G	L	T	M	D	Q	T	S	H	I	S	E	330
AAG	ATT	CCT	GGT	TTG	ACT	ATG	GAC	CAG	ACG	TCG	CAT	ATA	TCC	GAG	990
N	S	D	Y	V	T	T	T	T	T	T	T	T	T	T	336
TGG	TCG	GAT	TAT	GTA	ATT	TGA									1011

**Figure 6: Séquence nucléotidique et peptidique d'un gène de l'isoforme SAT 1' (L78443) d'*A. thaliana*.**

M L P V T S R R H F ATG TTG CCG GTC ACA AGT CGC CGC CAC TTC																30	10
T M S L Y M L R S S S P H I N																	25
ACA	ATG	TCC	CTA	TAT	ATG	CTC	CGT	TCA	TCT	TCT	CCA	CAC	ATC	AAT		75	75
H	H	S	F	L	L	P	S	F	V	S	S	K	F	K			40
CAT	CAC	TCT	TTC	CTT	CTT	CCT	TCT	TTT	GTT	TCC	TCC	AAA	TTC	AAA		120	55
H	H	T	L	S	P	P	P	S	P	P	P	P	P	P		165	70
CAC	CAT	ACT	TTA	TCT	CCT	CCT	CCT	TCT	CCT	CCT	CCT	CCT	CCT	CCT		210	35
M	A	A	C	I	D	T	C	R	T	G	K	P	Q	I		255	100
ATG	GCT	GCG	TGC	ATC	GAC	ACC	TGC	CGC	ACT	GGT	AAA	CCC	CAG	ATT		300	115
D	E	A	D	S	S	K	I	H	D	D	E	S	G	E		345	130
TCT	CCT	CGC	GAT	TCT	TCT	AAA	CAC	CAC	GAC	GAT	GAA	TCT	GGC	TTT		390	145
D	I	M	N	I	F	R	I	E	D	R	S	S	F	N		435	160
CGT	TAC	ATG	AAC	TAC	TTT	CGT	TAT	CCT	GAT	CGA	TCT	TCC	TTC	AAT		480	175
G	T	Q	T	K	T	L	H	T	R	P	L	L	E	D		525	190
GGA	ACC	CAG	ACC	AAA	ACC	CTC	CAT	ACT	CGT	CCT	TTG	CTT	GAA	GAT		570	205
L	D	R	D	A	E	V	D	D	V	W	A	K	I	R		615	220
CTC	GAT	CGC	GAC	GCT	GAA	GTC	GAT	GAT	GTT	TGG	GCC	AAA	ATC	CGA		660	235
L	C	A	K	S	D	I	A	K	L	P	I	V	S	A		705	250
GAA	GAG	GCT	AAA	TCT	GAT	ATC	GCC	AAA	GAA	CCT	ATT	GTT	TCC	GCT		750	265
I	V	H	A	S	I	V	S	Q	A	S	L	E	A	A		795	280
TAT	TAT	CAC	GCT	TCG	ATT	GTT	TCT	CAG	CGT	TCG	TTG	GAA	GCT	GCG		840	295
L	A	N	T	L	S	V	K	L	S	N	L	N	L	P		885	310
TTG	GCG	AAT	ACT	TTA	TCT	GTT	AAA	CTC	AGC	AAT	TTG	AAT	CTT	CCA		930	325
S	N	T	L	F	D	L	F	S	G	V	L	Q	G	N		975	340
AGC	AAC	ACG	CTT	TTC	GAT	TTG	TTC	TCT	GGT	GTT	CTT	CAA	GGA	AAC		1020	355
P	D	I	V	E	S	V	K	L	D	L	L	A	V	K		1065	370
CCA	GAT	ATT	GTT	GAA	TCT	GTC	AAG	CTA	GAT	CTT	TTA	GCT	GTT	AAG		1110	385
L	R	D	P	A	C	I	S	Y	V	H	C	F	L	H		1155	391
GAG	AGA	GAT	CCT	GCT	TGT	ATA	AGC	TAC	GTT	CAT	TGT	TTC	CTT	CAC			
F	K	G	F	L	A	C	Q	A	H	R	I	A	H	E			
TTT	AAA	GGC	TTC	CTC	GCT	TGT	CAA	GCG	CAT	CGT	ATT	GCT	CAT	GAG			
L	W	T	Q	D	R	K	L	L	A	L	L	I	Q	N			
CTT	TGG	ACT	CAG	GAC	AGA	AAA	ATC	CTA	GCT	TTG	TTG	ATC	CAG	AAC			
R	V	S	E	A	F	A	V	D	F	H	P	G	A	K			
AGA	GTC	TCT	GAA	GCC	TTC	GCT	GTT	GAT	TTT	CAC	CCT	GGA	GCT	AAA			
L	G	T	G	I	L	D	H	A	T	A	I	V	I				
ATC	GGT	ACC	GGG	ATT	TTG	CTA	GAC	CAT	GCT	ACG	GCT	ATT	GTG	ATC			
G	E	T	A	V	V	G	N	N	V	S	I	L	H	N			
GGT	GAG	ACG	GCG	GTT	GTG	GGG	AAC	AAT	GTT	TCG	ATT	CTC	CAT	AAC			
V	T	L	G	G	T	G	K	Q	C	G	D	R	H	P			
GTT	ACG	CTT	GGA	GGA	ACG	GGG	AAA	CAG	TGT	GGA	GAT	AGG	CAC	CCG			
K	L	G	D	G	V	L	I	G	A	G	T	C	I	L			
AAG	ATT	GGC	GAT	GGG	GTT	TTG	ATT	GGA	GCT	GGG	ACT	TGT	ATT	TTG			
G	N	I	T	I	G	E	G	A	K	I	G	A	G	S			
GGG	AAT	ATC	ACG	ATT	GGT	GAA	GGA	GCT	AAG	ATT	GGT	GCG	GGG	TCG			
L	V	L	K	D	V	P	P	R	T	T	A	V	G	N			
GTG	GTG	TTG	AAA	GAC	GTG	CCG	CCG	CGT	ACG	ACG	GCT	GTT	GGA	AAT			
E	A	R	L	L	G	G	K	D	N	P	K	T	H	D			
CCG	GCG	AGG	TTG	CTT	GGT	GGT	AAA	GAT	AAT	CCG	AAA	ACG	CAT	GAC			
L	-	V	G	-	T	M	-	Q	T	S	I	I	S	E			
AAG	ATT	CCT	GGT	TTG	ACT	ATG	GAC	CAG	ACG	TCG	CAT	ATA	TCC	GAG			
L	S	D	I	V	I	-	-	-	-	-	-	-	-	-			
TGG	TCG	GAT	TAT	GTA	ATT	TGA											

Figure 7 : Séquence nucléotidique et peptidique d'un gène de l'isoforme SAT 1 (U22964) d'*A. thaliana*.

M	V	D	L	S	S	F	S	L	L	F	A	F	S	V	S		
ATG	GTG	GAT	CTA	TCT	TCC	TTT	AGC	CTC	CTT	TTT	GCT	TTC	TCC	GTC	TCT	16	48
L	S	F	V	Q	S	K	R	V	C	D	S	S	L	S	S	32	
CTC	TCT	TTT	GTC	CAA	TCA	AAA	AGA	GTT	TGT	GAT	TCT	TCT	TTA	TCG	TCT		96
E	N	R	D	M	N	G	D	E	L	F	F	I	S	G	E	48	144
CCT	TGG	AGA	GAT	ATG	AAT	GGC	GAT	GAG	CTT	CCT	TTC	GAG	AGT	GGT	TTC	64	
E	V	Y	A	K	G	T	H	K	S	I	F	D	S	V	L	80	192
GAG	GTT	TAC	GCT	AAG	GGA	ACT	CAT	AAG	TCA	GAG	TTT	GAC	TCG	AAT	TTG		
L	D	P	R	S	D	P	I	N	D	A	I	R	E	E	A	80	240
CTT	GAT	CCT	CGT	TCT	GAT	CCT	ATT	TGG	GAT	GCT	ATA	AGA	GAA	GAA	GCT	96	288
N	L	E	A	E	K	E	F	L	L	S	F	L	Y	A		112	
AAA	CTT	GAG	GCA	GAG	AAA	GAG	CCT	ATT	TTG	AGT	AGC	TTC	TTG	TAT	GCT	128	336
G	I	L	A	H	D	C	L	E	Q	A	L	G	F	V	L	128	
GGT	ATC	TTA	GCA	CAT	GAT	TGT	TTA	GAG	CAA	GCT	TTA	GGG	TTT	GTT	CTA	144	432
A	N	R	L	Q	N	P	T	L	L	A	T	Q	I	L	D	160	
GCC	AAC	CGT	CTC	CAA	AAC	CCA	ACC	TTG	TTG	GCA	ACA	CAA	CTC	TTG	GAT	176	480
-	P	I	G	V	M	M	H	D	A	G	I	Q	S	S	I	192	
ATA	TTT	TAT	GGT	GTT	ATG	ATG	CAT	GAC	AAA	GGT	ATT	CAG	AGT	TCG	ATT	208	576
R	H	D	L	Q	A	F	K	D	R	C	P	A	C	L	S	224	
CGC	CAT	GAT	CTC	CAG	GCA	TTT	AAA	GAT	CGT	GAT	CCT	GCT	TGT	CTG	TCG	240	624
I	S	S	A	I	L	H	L	K	G	I	H	A	L	Q	A	256	
TAT	AGT	TCT	GCT	ATT	TTA	CAT	CTG	AAG	GGT	TAT	CAT	GCG	TTA	CAA	GCA	272	816
I	R	V	A	H	K	L	W	N	E	G	R	K	L	L	A	288	
TAT	AGG	GTT	GCG	GAT	AAA	CTG	TGG	AAT	GAA	GGG	AGG	AAA	CTA	TTA	GCT	304	912
L	A	L	Q	S	R	I	S	L	V	F	G	I	D	I	H	320	
CTT	GCA	TTG	CAA	AGC	CGA	ATA	AGC	GAG	GTT	TTT	GGC	ATT	GAC	ATA	CAT	336	1008
P	A	A	R	I	G	E	G	I	L	L	D	H	G	T	G	352	
CCA	GCG	GCA	AGA	ATT	GGG	GAG	GGA	ATA	TTG	TTG	GAT	CAT	GGA	ACT	GGA	368	1056
V	V	I	G	E	T	A	V	I	G	N	G	V	S	I	L	384	
GTG	GTC	ATT	GGT	GAG	ACC	GCT	GTG	ATA	GGC	AAC	GGT	GTC	TCG	ATC	TTA	400	1104
H	G	V	T	L	G	G	T	G	K	E	T	G	D	R	H	416	
CAT	GGT	GTG	ACT	TTA	GGA	GGA	ACC	GGA	AAG	GAA	ACT	GGC	GAT	CGC	CAC	432	
E	K	I	G	E	G	A	L	L	G	A	C	V	T	I	L	448	
CCA	AAG	ATA	GGT	GAA	GGT	GCA	TTG	CTT	GGA	GCT	TGT	GTG	ACT	ATA	CTT	464	
D	N	I	S	I	G	A	G	A	M	V	A	A	G	S	L	480	
GGT	AAC	ATA	AGC	ATA	GGT	GCT	GGA	GCA	ATG	GTA	GCT	GCA	GGT	TCA	CTT	496	
V	L	K	D	V	P	S	H	S	V	V	A	G	N	P	A	512	
GTG	TTA	AAA	GAC	GTT	CCT	TCG	CAT	AGT	GTG	GTG	GCT	GGA	AAT	CCT	GCA	528	
K	L	I	R	V	M	E	E	Q	D	P	S	L	A	M	K	544	
AAA	CTG	ATC	AGG	GTC	ATG	GAA	GAG	CAA	GAC	CCG	TCT	CTA	GCA	ATG	AAA	560	
H	D	A	T	K	E	F	F	R	H	V	A	D	G	Y	K	576	
CAC	GAT	GCT	ACT	AAA	GAG	TTC	TTT	CGA	CAT	GTA	GCT	GAT	GGT	TAC	AAA	592	
S	A	Q	S	N	G	P	S	L	S	A	G	D	T	E	K	608	
GGG	GCA	CAA	TCT	AAC	GGA	CCA	TCA	CTT	TCA	GCA	GGA	GAT	ACA	GAG	AAA	624	
G	H	T	N	S	T	S										640	
GGA	CAC	ACT	AAC	AGC	ACA	TCA	TGA									656	

**Figure 8: Sequence nucléotidique et peptidique du m RNA de la serine acetyltransferase SAT 2 putative chloroplastique d'*Arabidopsis thaliana* (L78444)**

M	A	C	I	N	G	E	N	R	D	F	S	S	S	S		
ATG	GCT	TGT	ATA	AAC	GGC	GAG	AAT	CGT	GAT	TTT	TCT	TCC	TCG	TCA	15	
S	L	S	S	L	P	M	I	V	S	R	N	F	S	A	30	45
TCT	TTG	TCT	TCT	CTT	CCA	ATG	ATT	GTC	TCC	CGG	AAC	TTT	TCT	GCC	45	90
R	D	D	G	E	T	G	D	E	F	P	F	E	R	I	135	
AGA	GAC	GAT	GGA	GAG	ACC	GGT	GAC	GAG	TTT	CCT	TTC	GAG	AGG	ATT	60	180
F	P	V	Y	A	R	G	T	L	N	P	V	A	D	P	75	
TTC	CCG	GTT	TAC	GCT	AGA	GGA	ACC	CTT	AAT	CCC	GTG	GCC	GAC	CCG	90	225
V	L	L	D	F	T	N	S	S	Y	D	P	I	W	D	105	
GTT	TTG	CTG	GAT	TTT	ACC	AAT	TCT	AGT	TAT	GAC	CCA	ATT	TGG	GAT	120	315
S	I	R	E	E	A	K	L	E	A	E	E	E	P	V	135	
TCT	ATA	AGA	GAA	GAA	GCT	AAG	CTT	GAG	GCA	GAA	GAG	GAG	CCG	GTT	150	405
L	S	S	F	L	Y	A	S	I	L	S	H	D	C	L	165	
TTG	AGT	AGC	TTC	TTG	TAT	GCT	AGT	ATC	TTG	TCG	CAT	GAC	TGT	TTA	180	540
E	Q	A	L	S	F	V	L	A	N	R	L	Q	N	P	195	
GAG	CAA	GCA	TTG	AGT	TTT	GTT	CTA	GCT	AAC	CGT	CTC	CAA	AAC	CCT	210	585
T	L	L	A	T	Q	L	M	D	I	F	C	N	V	M	225	
ACC	TTG	TTG	GCA	ACT	CAG	CTT	ATG	GAT	ATA	TTT	TGC	AAC	GTT	ATG	240	720
V	H	D	R	G	I	Q	S	S	I	R	L	D	V	Q	255	
GTA	CAT	GAC	AGA	GGT	ATT	CAA	AGC	TCG	ATT	CGT	CTT	GAT	GTT	CAG	270	810
A	F	K	D	R	D	P	A	C	L	S	Y	S	S	A	285	
GCA	TTC	AAA	GAC	AGA	GAT	CCT	GCT	TGT	CTA	TCG	TAT	AGT	TCG	GCT	299	900
I	L	H	L	K	G	Y	L	A	L	Q	A	Y	R	V		
ATT	TTA	CAT	CTG	AAG	GGC	TAT	CTT	GCA	CTG	CAG	GCG	TAT	AGA	GTA		
A	H	K	L	W	K	Q	G	R	K	L	L	A	L	A		
GCA	CAT	AAG	TTG	TGG	AAG	CAA	GGA	AGA	AAA	CTA	TTA	GCA	TTG	GCA		
L	Q	S	R	V	S	E	V	R	T	A	V	I	G	D		
CTG	CAA	AGC	CGA	GTA	AGC	GAG	GTA	AGA	ACT	GCT	GTG	ATA	GGC	GAC		
R	V	S	I	L	H	G	V	T	L	G	G	T	G	K		
CGT	GTC	TCA	ATT	TTG	CAT	GGT	GTG	ACA	TTA	GGA	GGA	ACT	GGG	AAA		
E	T	G	D	R	H	P	N	I	G	D	G	A	L	L		
GAA	ACC	GGT	GAC	CGC	CAT	CCA	AAT	ATA	GGC	GAC	GGT	GCT	CTT	CTT		
G	A	C	V	T	I	L	G	N	I	K	I	G	A	G		
GGA	GCA	TGT	GTG	ACT	ATA	CTT	GGT	AAC	ATT	AAG	ATA	GGC	GCT	GGA		
A	M	V	A	A	G	S	L	V	L	K	D	V	P	S		
GCA	ATG	GTA	GCT	GCT	GGT	TCG	CTT	GTG	TTA	AAG	GAT	GTT	CCT	TCG		
H	S	M	V	A	G	N	P	A	K	L	I	G	F	V		
CAT	AGC	ATG	GTG	GCT	GGA	AAT	CCA	GCA	AAA	CTC	ATC	GGG	TTT	GTT		
D	E	Q	D	P	S	M	T	M	E	H	G	E	S			
GAT	GAG	CAA	GAT	CCA	TCT	ATG	ACA	ATG	GAG	CAT	GGT	GAG	TCT	TGA		

**Figure 9: Sequence nucleotidique et en acides aminés du mRNA de la SAT4 putative chloroplastique d'*Arabidopsis thaliana*.**

8/12

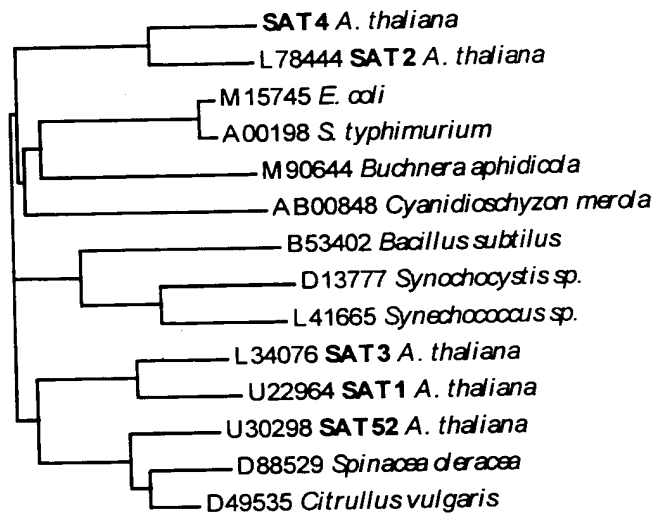


Figure 10 Dendrogramme des serine acétyltransferase issues de plusieurs organismes.

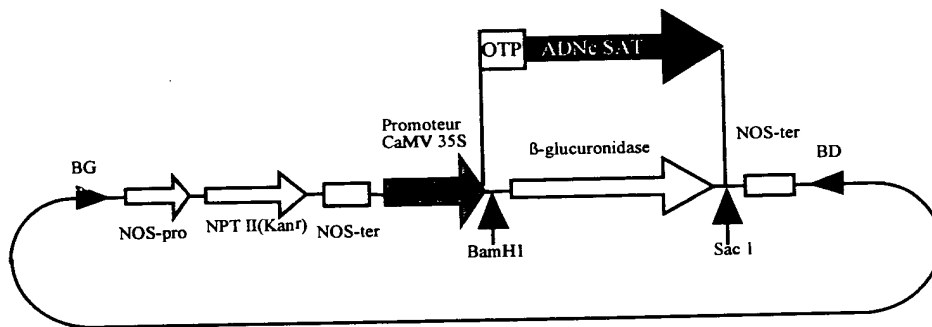


Figure 11: Procédure de clonage de l'OTP/Serine acétyltransférase SAT3 ou SAT (insensible à la cystéine, par exemple SAT1 tronqué) dans le vecteur pBI121.

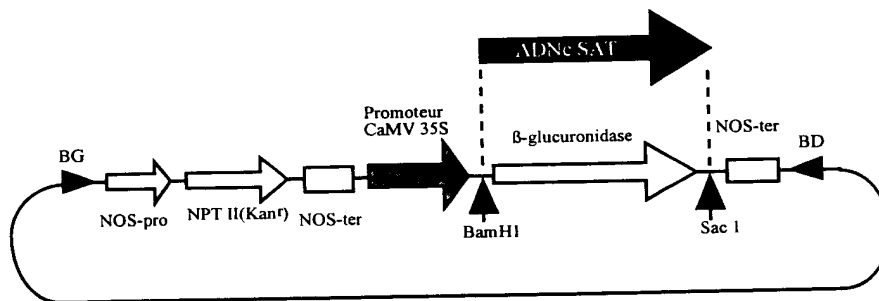


Figure 12: Procédure de clonage de la Serine acétyltransférase SAT1' ; SAT1 ; SAT2 ; SAT3, SAT3' : SAT4, ou toutes SATs dans le vecteur pBI121.

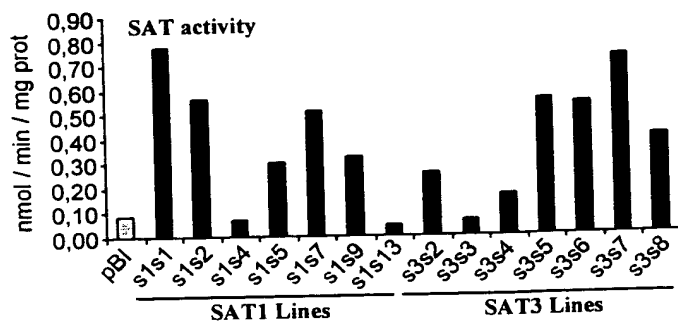


Figure 13

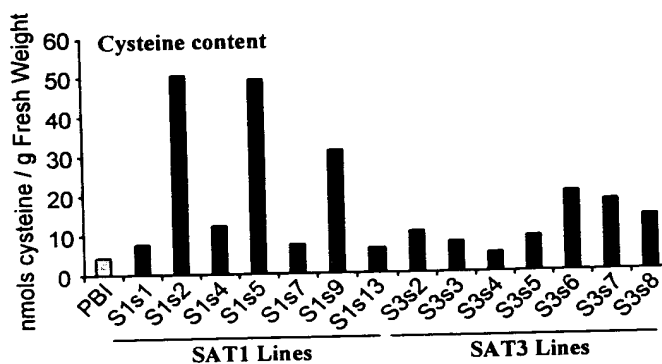


Figure 14

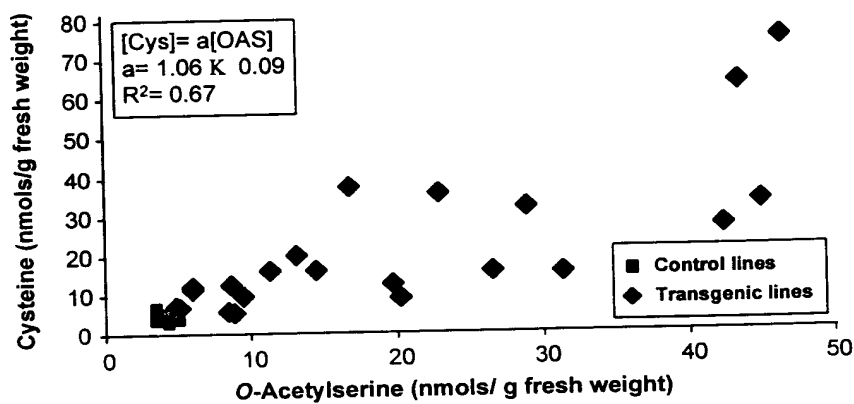


Figure 15

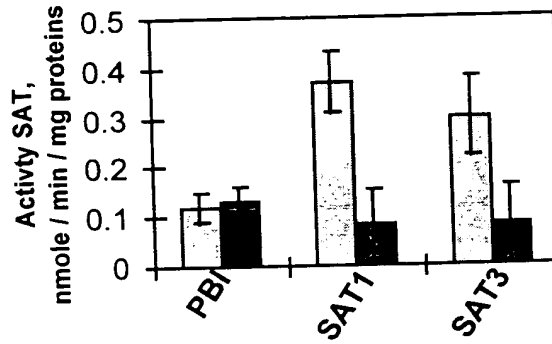


Figure 16

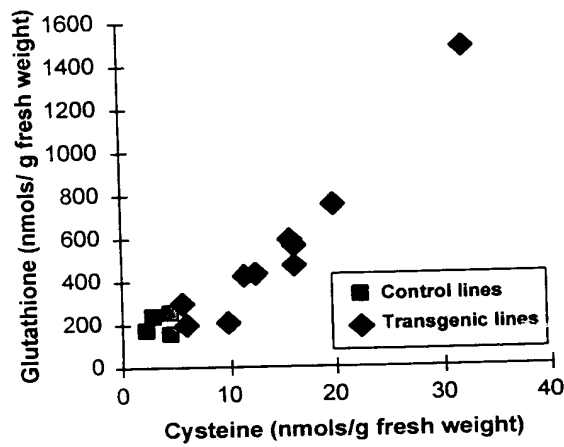
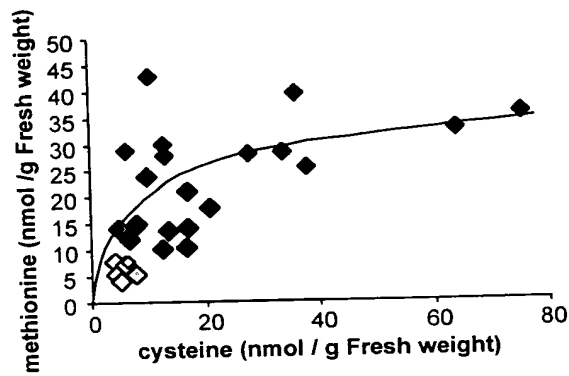


Figure 17



11/12

09/486334

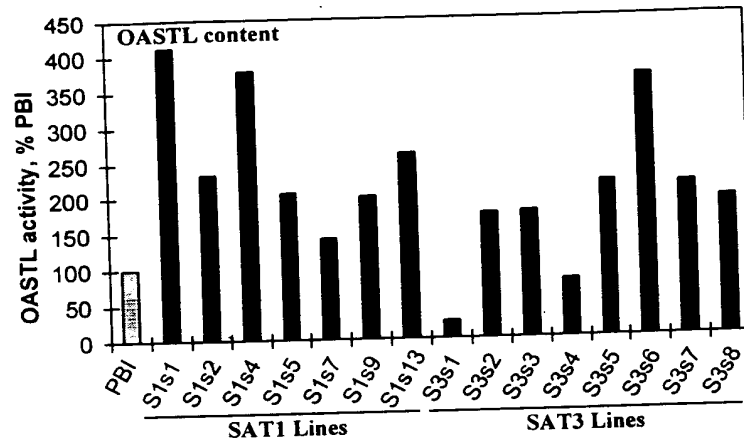


Figure 19

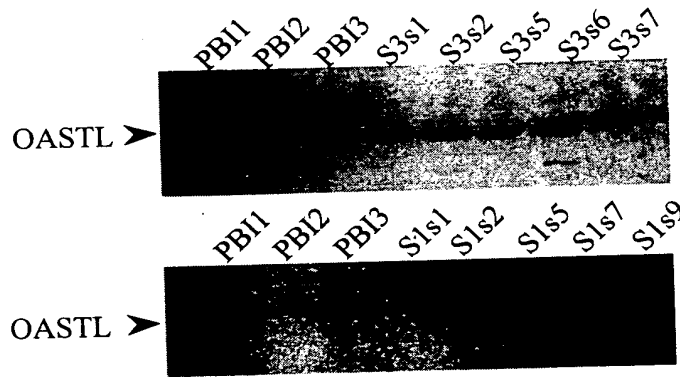


Figure 20

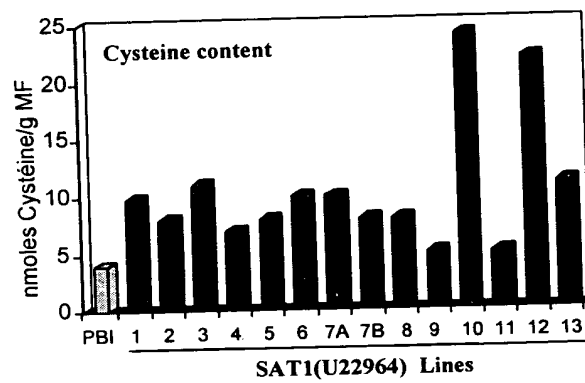


Figure 21

12/12

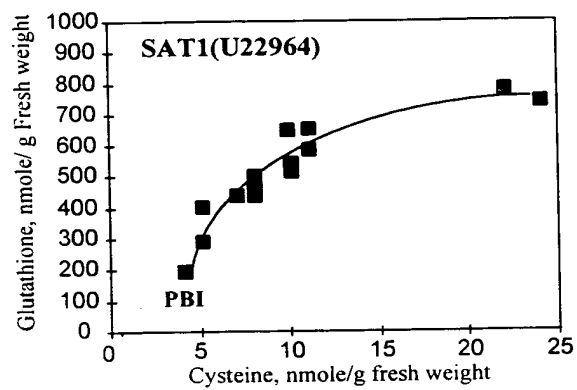


Figure 22

## LISTE DE SEQUENCES

09/486334  
514 Rec'd PCT/PTO 22 FEB 2000

&lt;110&gt; RHONE-POULENC AGRO

<120> procédé pour augmenter la teneur en cystéine,  
méthionine et glutathion chez les plantes et plantes  
obtenues

&lt;130&gt;

&lt;140&gt;

&lt;141&gt;

&lt;150&gt; FR9816163

&lt;151&gt; 1998-12-17

&lt;160&gt; 17

&lt;170&gt; PatentIn Ver. 2.1

&lt;210&gt; 1

&lt;211&gt; 984

&lt;212&gt; ADN

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (31)..(972)

&lt;400&gt; 1

gagagaggat cctcttttcca atcataaacc atg gca aca tgc ata gac aca tgc	54
Met Ala Thr Cys Ile Asp Thr Cys	
1 5	
cga acc ggt aat acc caa gac gat gat tcc cgg ttc tgt tgc atc aag	102
Arg Thr Gly Asn Thr Gln Asp Asp Asp Ser Arg Phe Cys Cys Ile Lys	
10 15 20	
aat ttc ttt cga ccc ggt ttc tct gta aac cgg aag att cac cac acc	150
Asn Phe Phe Arg Pro Gly Phe Ser Val Asn Arg Lys Ile His His Thr	
25 30 35 40	
caa atc gaa gat gac gat gat gtc tgg atc aag atg ctt gaa gaa gcc	198
Gln Ile Glu Asp Asp Asp Asp Val Trp Ile Lys Met Leu Glu Glu Ala	
45 50 55	
aaa tcc gat gtt aaa caa gaa ccc att tta tca aac tac tac tac gct	246
Lys Ser Asp Val Lys Gln Glu Pro Ile Leu Ser Asn Tyr Tyr Tyr Ala	
60 65 70	
tcg atc aca tct cat cga tct tta gag tct gct tta gct cac atc ctc	294
Ser Ile Thr Ser His Arg Ser Leu Glu Ser Ala Leu Ala His Ile Leu	
75 80 85	
tcc gta aag ctc agc aat tta aac cta cca agc aac aca ctc ttc gaa	342
Ser Val Lys Leu Ser Asn Leu Asn Leu Pro Ser Asn Thr Leu Phe Glu	
90 95 100	
ctg ttc ata agc gtt tta gaa gaa agc cct gag atc atc gaa tcc acg	390
Leu Phe Ile Ser Val Leu Glu Glu Ser Pro Glu Ile Ile Glu Ser Thr	
105 110 115 120	
aag caa gat ctt ata gca gtc aaa gaa aga gac cca gct tgt ata agc	438
Lys Gln Asp Leu Ile Ala Val Lys Glu Arg Asp Pro Ala Cys Ile Ser	
125 130 135	

tac gtt cat tgc ttc ttg ggc ttc aaa ggc ttc ctc gct tgt caa gct	486
Tyr Val His Cys Phe Leu Gly Phe Lys Gly Phe Leu Ala Cys Gln Ala	
140 145 150	
cat cga ata gct cat acc ctc tgg aaa cag aac aga aaa atc gta gct	534
His Arg Ile Ala His Thr Leu Trp Lys Gln Asn Arg Lys Ile Val Ala	
155 160 165	
tta ttg atc caa aac aga gta tca gaa tct ttc gcc gtc gat att cat	582
Leu Leu Ile Gln Asn Arg Val Ser Glu Ser Phe Ala Val Asp Ile His	
170 175 180	
ccc gga gcg aag atc gga aaa ggg att ctt tta gac cat gcg acg ggc	630
Pro Gly Ala Lys Ile Gly Lys Gly Ile Leu Leu Asp His Ala Thr Gly	
185 190 195 200	
gtg gtg atc gga gag acg gcg gtg gtt gga gac aat gtt tcg att cta	678
Val Val Ile Gly Glu Thr Ala Val Val Gly Asp Asn Val Ser Ile Leu	
205 210 215	
cac gga gtg acc ttg gga gga aca ggg aaa cag agt ggt gat cgg cat	726
His Gly Val Thr Leu Gly Gly Thr Gly Lys Gln Ser Gly Asp Arg His	
220 225 230	
ccg aag att ggt gat ggt gtg ttg att gga gct ggg agt tgt ata ttg	774
Pro Lys Ile Gly Asp Gly Val Leu Ile Gly Ala Gly Ser Cys Ile Leu	
235 240 245	
ggg aat ata aca atc ggt gag gga gct aag att gga tca ggg tcg gtg	822
Gly Asn Ile Thr Ile Gly Glu Gly Ala Lys Ile Gly Ser Gly Ser Val	
250 255 260	
gtg gtt aag gat gtg ccg gcg cgt acg acg gcg gtt gga aat ccg gcg	870
Val Val Lys Asp Val Pro Ala Arg Thr Thr Ala Val Gly Asn Pro Ala	
265 270 275 280	
agg ttg att ggt ggg aaa gag aat ccg aga aaa cat gat aag att cct	918
Arg Leu Ile Gly Gly Lys Glu Asn Pro Arg Lys His Asp Lys Ile Pro	
285 290 295	
tgt ctg act atg gac cag aca tcg tat tta acc gag tgg tct gat tat	966
Cys Leu Thr Met Asp Gln Thr Ser Tyr Leu Thr Glu Trp Ser Asp Tyr	
300 305 310	
gtg att taacacaaat gt	984
Val Ile	

<210> 2  
 <211> 974  
 <212> ADN  
 <213> Arabidopsis thaliana

<220>  
 <221> CDS  
 <222> (31)..(966)

<400> 2	
gagagaggat cctcttatcg ccgcgttaat atg cca ccg gcc gga gaa ctc cga	54
Met Pro Pro Ala Gly Glu Leu Arg	
1 5	

cat caa tct cca tca aag gag aaa cta tct tcc gtt acc caa tcc gat	102
His Gln Ser Pro Ser Lys Glu Lys Leu Ser Ser Val Thr Gln Ser Asp	
10 15 20	

gaa gca gaa gca gcg tca gca gcg ata tct gcg gca gct gca gat gcg Glu Ala Glu Ala Ala Ser Ala Ala Ile Ser Ala Ala Ala Asp Ala 25 30 35 40	150
gaa gct gcc gga tta tgg aca cag atc aag gcg gaa gct cgc cgt gat Glu Ala Ala Gly Leu Trp Thr Gln Ile Lys Ala Glu Ala Arg Arg Asp 45 50 55	198
gct gag gcg gag cca gct tta gct agc tat cta tat tcg acg att ctt Ala Glu Ala Glu Pro Ala Leu Ala Ser Tyr Leu Tyr Ser Thr Ile Leu 60 65 70	246
tct cat tcg tct ctt gaa cga tct atc tcg ttt cat cta gga aac aag Ser His Ser Ser Leu Glu Arg Ser Ile Ser Phe His Leu Gly Asn Lys 75 80 85	294
ctt tgt tcc tca acg ctt tta tcc aca ctt tta tac gat ctg ttc tta Leu Cys Ser Ser Thr Leu Leu Ser Thr Leu Leu Tyr Asp Leu Phe Leu 90 95 100	342
aac act ttt tcc tcc gat cct tct ctt cgt aac gcc acc gtc gca gat Asn Thr Phe Ser Ser Asp Pro Ser Leu Arg Asn Ala Thr Val Ala Asp 105 110 115 120	390
cta cgc gct gct cgt gtt cgt gat cct gct tgt atc tcg ttc tct cat Leu Arg Ala Ala Arg Val Arg Asp Pro Ala Cys Ile Ser Phe Ser His 125 130 135	438
tgt ctc ctc aat tac aaa ggc ttc tta gct att cag gcg cat cgt gta Cys Leu Leu Asn Tyr Lys Gly Phe Leu Ala Ile Gln Ala His Arg Val 140 145 150	486
tca cac aag cta tgg aca caa tca cgg aag cca tta gca tta gct cta Ser His Lys Leu Trp Thr Gln Ser Arg Lys Pro Leu Ala Leu Ala Leu 155 160 165	534
cac tca aga atc tcc gat gta ttc gct gtt gat atc cat cca gca gcg His Ser Arg Ile Ser Asp Val Phe Ala Val Asp Ile His Pro Ala Ala 170 175 180	582
aag atc gga aaa ggg ata ctt cta gac cac gca acc gga gtt gta gtc Lys Ile Gly Lys Gly Ile Leu Leu Asp His Ala Thr Gly Val Val Val 185 190 195 200	630
gga gaa aca gcg gtg att ggg aac aat gtt tca atc ctt cac cat gtg Gly Glu Thr Ala Val Ile Gly Asn Asn Val Ser Ile Leu His His Val 205 210 215	678
aca cta ggt gga aca ggt aaa gct tgt gga gat aga cat ccg aag atc Thr Leu Gly Gly Thr Gly Lys Ala Cys Gly Asp Arg His Pro Lys Ile 220 225 230	726
ggt gac ggt tgt ttg att gga gct gga gcg act att ctt gga aat gtg Gly Asp Gly Cys Leu Ile Gly Ala Gly Ala Thr Ile Leu Gly Asn Val 235 240 245	774
aag att ggt gca ggt gct aaa gta gga gct ggt tct gtt gtg ctg att Lys Ile Gly Ala Gly Ala Lys Val Gly Ala Gly Ser Val Val Leu Ile 250 255 260	822
gac gtg cct tgt cga ggt act gcg gtt ggg aat ccg gcg aga ctt gtc Asp Val Pro Cys Arg Gly Thr Ala Val Gly Asn Pro Ala Arg Leu Val 265 270 275 280	870

gga ggg aaa gag aag cca acg att cat gat gag gaa tgt cct gga gaa 918  
 Gly Gly Lys Glu Lys Pro Thr Ile His Asp Glu Glu Cys Pro Gly Glu 295  
 285

tcg atg gat cat act tca ttc atc tcg gaa tgg tca gat tac atc ata 966  
 Ser Met Asp His Thr Ser Phe Ile Ser Glu Trp Ser Asp Tyr Ile Ile 310  
 300

taaagttg 974

<210> 3  
 <211> 1048  
 <212> ADN  
 <213> Arabidopsis thaliana

<220>  
 <221> CDS  
 <222> (31)..(1038)

<400> 3  
 gagagaggat cccctcctcc tctcctcct atg gct gcg tgc atc gac acc tgc 54  
 Met Ala Ala Cys Ile Asp Thr Cys 5  
 1

cgc act ggt aaa ccc cag att tct cct cgc gat tct tct aaa cac cac 102  
 Arg Thr Gly Lys Pro Gln Ile Ser Pro Arg Asp Ser Ser Lys His His 20  
 10

gac gat gaa tct ggc ttt cgt tac atg aac tac ttc cgt tat cct gat 150  
 Asp Asp Glu Ser Gly Phe Arg Tyr Met Asn Tyr Phe Arg Tyr Pro Asp 40  
 25 30

cga tct tcc ttc aat gga acc cag acc aaa acc ctc cat act cgt cct 198  
 Arg Ser Ser Phe Asn Gly Thr Gln Thr Lys Thr Leu His Thr Arg Pro 55  
 45 50

ttg ctt gaa gat ctc gat cgc gac gct gaa gtc gat gat gtt tgg gcc 246  
 Leu Leu Glu Asp Leu Asp Arg Asp Ala Glu Val Asp Asp Val Trp Ala 70  
 60 65

aaa atc cga gaa gag gct aaa tct gat atc gcc aaa gaa cct att gtt 294  
 Lys Ile Arg Glu Glu Ala Lys Ser Asp Ile Ala Lys Glu Pro Ile Val 85  
 75 80

tcc gct tat tat cac gct tcg att gtt tct cag cgt tcg ttg gaa gct 342  
 Ser Ala Tyr Tyr His Ala Ser Ile Val Ser Gln Arg Ser Leu Glu Ala 100  
 90 95

gcg ttg gcg aat act tta tct gtt aaa ctc agc aat ttg aat ctt cca 390  
 Ala Leu Ala Asn Thr Leu Ser Val Lys Leu Ser Asn Leu Asn Leu Pro 120  
 105 110 115

agc aac acg ctt ttc gat ttg ttc tct ggt gtt ctt caa gga aac cca 438  
 Ser Asn Thr Leu Phe Asp Leu Phe Ser Gly Val Leu Gln Gly Asn Pro 135  
 125 130

gat att gtt gaa tct gtc aag cta gat ctt tta gct gtt aag gag aga 486  
 Asp Ile Val Glu Ser Val Lys Leu Asp Leu Leu Ala Val Lys Glu Arg 150  
 140 145

gat cct gct tgt ata agc tac gtt cat tgt ttc ctt cac ttt aaa ggc 534  
 Asp Pro Ala Cys Ile Ser Tyr Val His Cys Phe Leu His Phe Lys Gly 165  
 155 160

ttc ctc gct tgt caa gcg cat cgt att gct cat gag ctt tgg act cag 582  
 Phe Leu Ala Cys Gln Ala His Arg Ile Ala His Glu Leu Trp Thr Gln  
 170 175 180

gac aga aaa atc cta gct ttg ttg atc cag aac aga gtc tct gaa gcc 630  
 Asp Arg Lys Ile Leu Ala Leu Leu Ile Gln Asn Arg Val Ser Glu Ala  
 185 190 195 200

ttc gct gtt gat ttc cac cct gga gct aaa atc ggt acc ggg att ttg 678  
 Phe Ala Val Asp Phe His Pro Gly Ala Lys Ile Gly Thr Gly Ile Leu  
 205 210 215

cta gac cat gct acg gct att gtg atc ggt gag acg gcg gtt gtg ggg 726  
 Leu Asp His Ala Thr Ala Ile Val Ile Gly Glu Thr Ala Val Val Gly  
 220 225 230

aac aat gtt tcg att ctc cat aac gtt acg ctt gga gga acg ggg aaa 774  
 Asn Asn Val Ser Ile Leu His Asn Val Thr Leu Gly Gly Thr Gly Lys  
 235 240 245

cag tgt gga gat agg cac ccg aag att ggc gat ggg gtt ttg att gga 822  
 Gln Cys Gly Asp Arg His Pro Lys Ile Gly Asp Gly Val Leu Ile Gly  
 250 255 260

gct ggg act tgt att ttg ggg aat atc acg att ggt gaa gga gct aag 870  
 Ala Gly Thr Cys Ile Leu Gly Asn Ile Thr Ile Gly Glu Gly Ala Lys  
 265 270 275 280

att ggt gcg ggg tcg gtg gtg ttg aaa gac gtg ccg ccg cgt acg acg 918  
 Ile Gly Ala Gly Ser Val Val Leu Lys Asp Val Pro Pro Arg Thr Thr  
 285 290 295

gct gtt gga aat ccg gcg agg ttg ctt ggt ggt aaa gat aat ccg aaa 966  
 Ala Val Gly Asn Pro Ala Arg Leu Leu Gly Gly Lys Asp Asn Pro Lys  
 300 305 310

acg cat gac aag att cct ggt ttg act atg gac cag acg tcg cat ata 1014  
 Thr His Asp Lys Ile Pro Gly Leu Thr Met Asp Gln Thr Ser His Ile  
 315 320 325

tcc gag tgg tcg gat tat gta att tgaaaaagtc 1048  
 Ser Glu Trp Ser Asp Tyr Val Ile  
 330 335

<210> 4  
 <211> 1213  
 <212> ADN  
 <213> Arabidopsis thaliana

<220>  
 <221> CDS  
 <222> (31)..(1203)

<220>  
 <221> sig\_peptide  
 <222> (31)..(219)

<400> 4  
 gagagaggat ccggccgaga aaaaaaaaaa atg ttg ccg gtc aca agt cgc cgc 54  
 Met Leu Pro Val Thr Ser Arg Arg  
 1 5

cac ttc aca atg tcc cta tat atg ctc cgt tca tct tct cca cac atc 102  
 His Phe Thr Met Ser Leu Tyr Met Leu Arg Ser Ser Ser Pro His Ile  
 10 15 20

aat cat cac tct ttc ctt ctt cct tct ttt gtt tcc tcc aaa ttc aaa	150
Asn His His Ser Phe Leu Leu Pro Ser Phe Val Ser Ser Lys Phe Lys	
25 30 35 40	
cac cat act tta tct cct cct cct tct cct cct cct cct cct atg	198
His His Thr Leu Ser Pro Pro Pro Ser Pro Pro Pro Pro Pro Met	
45 50 55	
gct gcg tgc atc gac acc tgc cgc act ggt aaa ccc cag att tct cct	246
Ala Ala Cys Ile Asp Thr Cys Arg Thr Gly Lys Pro Gln Ile Ser Pro	
60 65 70	
cgc gat tct tct aaa cac cac gac gat gaa tct ggc ttt cgt tac atg	294
Arg Asp Ser Ser Lys His His Asp Asp Glu Ser Gly Phe Arg Tyr Met	
75 80 85	
aac tac ttc cgt tat cct gat cga tct tcc ttc aat gga acc cag acc	342
Asn Tyr Phe Arg Tyr Pro Asp Arg Ser Ser Phe Asn Gly Thr Gln Thr	
90 95 100	
aaa acc ctc cat act cgt cct ttg ctt gaa gat ctc gat cgc gac gct	390
Lys Thr Leu His Thr Arg Pro Leu Leu Glu Asp Leu Asp Arg Asp Ala	
105 110 115 120	
gaa gtc gat gat gtt tgg gcc aaa atc cga gaa gag gct aaa tct gat	438
Glu Val Asp Asp Val Trp Ala Lys Ile Arg Glu Glu Ala Lys Ser Asp	
125 130 135	
atc gcc aaa gaa cct att gtt tcc gct tat tat cac gct tcg att gtt	486
Ile Ala Lys Glu Pro Ile Val Ser Ala Tyr Tyr His Ala Ser Ile Val	
140 145 150	
tct cag cgt tcg ttg gaa gct gcg ttg gcg aat act tta tct gtt aaa	534
Ser Gln Arg Ser Leu Glu Ala Ala Leu Ala Asn Thr Leu Ser Val Lys	
155 160 165	
ctc agc aat ttg aat ctt cca agc aac acg ctt ttc gat ttg ttc tct	582
Leu Ser Asn Leu Asn Leu Pro Ser Asn Thr Leu Phe Asp Leu Phe Ser	
170 175 180	
ggt gtt ctt caa gga aac cca gat att gtt gaa tct gtc aag cta gat	630
Gly Val Leu Gln Gly Asn Pro Asp Ile Val Glu Ser Val Lys Leu Asp	
185 190 195 200	
ctt tta gct gtt aag gag aga gat cct gct tgt ata agc tac gtt cat	678
Leu Leu Ala Val Lys Glu Arg Asp Pro Ala Cys Ile Ser Tyr Val His	
205 210 215	
tgt ttc ctt cac ttt aaa ggc ttc ctc gct tgt caa gcg cat cgt att	726
Cys Phe Leu His Phe Lys Gly Phe Leu Ala Cys Gln Ala His Arg Ile	
220 225 230	
gct cat gag ctt tgg act cag gac aga aaa atc cta gct ttg ttg atc	774
Ala His Glu Leu Trp Thr Gln Asp Arg Lys Ile Leu Ala Leu Leu Ile	
235 240 245	
cag aac aga gtc tct gaa gcc ttc gct gtt gat ttc cac cct gga gct	822
Gln Asn Arg Val Ser Glu Ala Phe Ala Val Asp Phe His Pro Gly Ala	
250 255 260	
aaa atc ggt acc ggg att ttg cta gac cat gct acg gct att gtg atc	870
Lys Ile Gly Thr Gly Ile Leu Leu Asp His Ala Thr Ala Ile Val Ile	
265 270 275 280	

ggt gag acg gcg gtt gtg ggg aac aat gtt tcg att ctc cat aac gtt 918  
 Gly Glu Thr Ala Val Val Gly Asn Asn Val Ser Ile Leu His Asn Val 295  
 285 290

acg ctt gga gga acg ggg aaa cag tgt gga gat agg cac ccg aag att 966  
 Thr Leu Gly Gly Thr Gly Lys Gln Cys Gly Asp Arg His Pro Lys Ile 310  
 300 305

ggc gat ggg gtt ttg att gga gct ggg act tgt att ttg ggg aat atc 1014  
 Gly Asp Gly Val Leu Ile Gly Ala Gly Thr Cys Ile Leu Gly Asn Ile 325  
 315 320

acg att ggt gaa gga gct aag att ggt gcg ggg tcg gtg gtg ttg aaa 1062  
 Thr Ile Gly Glu Gly Ala Lys Ile Gly Ala Gly Ser Val Val Leu Lys 340  
 330 335

gac gtg ccg ccg cgt acg acg gct gtt gga aat ccg gcg agg ttg ctt 1110  
 Asp Val Pro Pro Arg Thr Thr Ala Val Gly Asn Pro Ala Arg Leu Leu 360  
 345 350 355

ggt ggt aaa gat aat ccg aaa acg cat gac aag att cct ggt ttg act 1158  
 Gly Gly Lys Asp Asn Pro Lys Thr His Asp Lys Ile Pro Gly Leu Thr 375  
 365 370

atg gac cag acg tcg cat ata tcc gag tgg tcg gat tat gta att 1203  
 Met Asp Gln Thr Ser His Ile Ser Glu Trp Ser Asp Tyr Val Ile 390  
 380 385

tgaaaaagtc 1213

<210> 5  
 <211> 1080  
 <212> ADN  
 <213> Arabidopsis thaliana

<220>  
 <221> CDS  
 <222> (1)..(1080)

<220>  
 <221> transit\_peptide  
 <222> (1)..(96)

<400> 5  
 atg gtg gat cta tct tcc ttt agc ctc ctt ttt gct ttc tcc gtc tct 48  
 Met Val Asp Leu Ser Ser Phe Ser Leu Leu Phe Ala Phe Ser Val Ser 15  
 1 5 10

ctc tct ttt gtc caa tca aaa aga gtt tgt gat tct tct tta tcg tct 96  
 Leu Ser Phe Val Gln Ser Lys Arg Val Cys Asp Ser Ser Leu Ser Ser 30  
 20 25

cct tgg aga gat atg aat ggc gat gag ctt cct ttc gag agt ggt ttc 144  
 Pro Trp Arg Asp Met Asn Gly Asp Glu Leu Pro Phe Glu Ser Gly Phe 45  
 35 40

gag gtt tac gct aag gga act cat aag tca gag ttt gac tcg aat ttg 192  
 Glu Val Tyr Ala Lys Gly Thr His Lys Ser Glu Phe Asp Ser Asn Leu 60  
 50 55

ctt gat cct cgt tct gat cct att tgg gat gct ata aga gaa gaa gct 240  
 Leu Asp Pro Arg Ser Asp Pro Ile Trp Asp Ala Ile Arg Glu Glu Ala 80  
 65 70 75

aaa ctt gag gca gag aaa gag cct att ttg agt agc ttc ttg tat gct Lys Leu Glu Ala Glu Lys Glu Pro Ile Leu Ser Ser Phe Leu Tyr Ala 85 90 95	288
ggt atc tta gca cat gat tgt tta gag caa gct tta ggg ttt gtt cta Gly Ile Leu Ala His Asp Cys Leu Glu Gln Ala Leu Gly Phe Val Leu 100 105 110	336
gcc aac cgt ctc caa aac cca acc ttg ttg gca aca caa ctc ttg gat Ala Asn Arg Leu Gln Asn Pro Thr Leu Leu Ala Thr Gln Leu Leu Asp 115 120 125	384
ata ttt tat ggt gtt atg atg cat gac aaa ggt att cag agt tcg att Ile Phe Tyr Gly Val Met Met His Asp Lys Gly Ile Gln Ser Ser Ile 130 135 140	432
cgc cat gat ctc cag gca ttt aaa gat cgt gat cct gct tgt ctg tcg Arg His Asp Leu Gln Ala Phe Lys Asp Arg Asp Pro Ala Cys Leu Ser 145 150 155 160	480
tat agt tct gct att tta cat ctg aag ggt tat cat gcg tta caa gca Tyr Ser Ser Ala Ile Leu His Leu Lys Gly Tyr His Ala Leu Gln Ala 165 170 175	528
tat agg gtt gcg cat aaa ctg tgg aat gaa ggg agg aaa cta tta gct Tyr Arg Val Ala His Lys Leu Trp Asn Glu Gly Arg Lys Leu Leu Ala 180 185 190	576
ctt gca ttg caa agc cga ata agc gag gtt ttt ggc att gac ata cat Leu Ala Leu Gln Ser Arg Ile Ser Glu Val Phe Gly Ile Asp Ile His 195 200 205	624
cca gcg gca aga att ggg gag gga ata ttg ttg gat cat gga act gga Pro Ala Ala Arg Ile Gly Glu Gly Ile Leu Leu Asp His Gly Thr Gly 210 215 220	672
gtg gtc att ggt gag acc gct gtg ata ggc aac ggt gtc tcg atc tta Val Val Ile Gly Glu Thr Ala Val Ile Gly Asn Gly Val Ser Ile Leu 225 230 235 240	720
cat ggt gtg act tta gga gga acc gga aag gaa act ggc gat cgc cac His Gly Val Thr Leu Gly Gly Thr Gly Lys Glu Thr Gly Asp Arg His 245 250 255	768
cca aag ata ggt gaa ggt gca ttg ctt gga gct tgt gtg act ata ctt Pro Lys Ile Gly Glu Gly Ala Leu Leu Gly Ala Cys Val Thr Ile Leu 260 265 270	816
ggt aac ata agc ata ggt gct gga gca atg gta gct gca ggt tca ctt Gly Asn Ile Ser Ile Gly Ala Gly Ala Met Val Ala Ala Gly Ser Leu 275 280 285	864
gtg tta aaa gac gtt cct tcg cat agt gtg gtg gct gga aat cct gca Val Leu Lys Asp Val Pro Ser His Ser Val Val Ala Gly Asn Pro Ala 290 295 300	912
aaa ctg atc agg gtc atg gaa gag caa gac ccg tct cta gca atg aaa Lys Leu Ile Arg Val Met Glu Glu Gln Asp Pro Ser Leu Ala Met Lys 305 310 315 320	960
cac gat gct act aaa gag ttc ttt cga cat gta gct gat ggt tac aaa His Asp Ala Thr Lys Glu Phe Phe Arg His Val Ala Asp Gly Tyr Lys 325 330 335	1008

ggg gca caa tct aac gga cca tca ctt tca gca gga gat aca gag aaa 1056  
 Gly Ala Gln Ser Asn Gly Pro Ser Leu Ser Ala Gly Asp Thr Glu Lys  
 340 345 350

gga cac act aac agc aca tca tga 1080  
 Gly His Thr Asn Ser Thr Ser  
 355 360

<210> 6  
 <211> 900  
 <212> ADN  
 <213> Arabidopsis thaliana

<220>  
 <221> CDS  
 <222> (1)..(900)

<220>  
 <221> transit\_peptide  
 <222> (1)..(90)

<400> 11  
 atg gct tgt ata aac ggc gag aat cgt gat ttt tct tcc tcg tca tct 48  
 Met Ala Cys Ile Asn Gly Glu Asn Arg Asp Phe Ser Ser Ser Ser Ser  
 1 5 10 15  
 ttg tct tct ctt cca atg att gtc tcc cgg aac ttt tct gcc aga gac 96  
 Leu Ser Ser Leu Pro Met Ile Val Ser Arg Asn Phe Ser Ala Arg Asp  
 20 25 30  
 gat gga gag acc ggt gac gag ttt cct ttc gag agg att ttc ccg gtt 144  
 Asp Gly Glu Thr Gly Asp Glu Phe Pro Phe Glu Arg Ile Phe Pro Val  
 35 40 45  
 tac gct aga gga acc ctt aat ccc gtg gcc gac ccg gtt ttg ctg gat 192  
 Tyr Ala Arg Gly Thr Leu Asn Pro Val Ala Asp Pro Val Leu Leu Asp  
 50 55 60  
 ttt acc aat tct agt tat gac cca att tgg gat tct ata aga gaa gaa 240  
 Phe Thr Asn Ser Ser Tyr Asp Pro Ile Trp Asp Ser Ile Arg Glu Glu  
 65 70 75 80  
 gct aag ctt gag gca gaa gag gag ccg gtt ttg agt agc ttc ttg tat 288  
 Ala Lys Leu Glu Ala Glu Glu Glu Pro Val Leu Ser Ser Phe Leu Tyr  
 85 90 95  
 gct agt atc ttg tcg cat gac tgt tta gag caa gca ttg agt ttt gtt 336  
 Ala Ser Ile Leu Ser His Asp Cys Leu Glu Gln Ala Leu Ser Phe Val  
 100 105 110  
 cta gct aac cgt ctc caa aac cct acc ttg ttg gca act cag ctt atg 384  
 Leu Ala Asn Arg Leu Gln Asn Pro Thr Leu Leu Ala Thr Gln Leu Met  
 115 120 125  
 gat ata ttt tgc aac gtt atg gta cat gac aga ggt att caa agc tcg 432  
 Asp Ile Phe Cys Asn Val Met Val His Asp Arg Gly Ile Gln Ser Ser  
 130 135 140  
 att cgt ctt gat gtt cag gca ttc aaa gac aga gat cct gct tgt cta 480  
 Ile Arg Leu Asp Val Gln Ala Phe Lys Asp Arg Asp Pro Ala Cys Leu  
 145 150 155 160  
 tcg tat agt tcg gct att tta cat ctg aag ggc tat ctt gca ctg cag 528  
 Ser Tyr Ser Ser Ala Ile Leu His Leu Lys Gly Tyr Leu Ala Leu Gln  
 165 170 175

gcg tat aga gta gca cat aag ttg tgg aag caa gga aga aaa cta tta 576  
 Ala Tyr Arg Val Ala His Lys Leu Trp Lys Gln Gly Arg Lys Leu Leu  
 180 185 190

gca ttg gca ctg caa agc cga gta agc gag gta aga act gct gtg ata 624  
 Ala Leu Ala Leu Gln Ser Arg Val Ser Glu Val Arg Thr Ala Val Ile  
 195 200 205

ggc gac cgt gtc tca att ttg cat ggt gtg aca tta gga gga act ggg 672  
 Gly Asp Arg Val Ser Ile Leu His Gly Val Thr Leu Gly Gly Thr Gly  
 210 215 220

aaa gaa acc ggt gac cgc cat cca aat ata ggc gac ggt gct ctt ctt 720  
 Lys Glu Thr Gly Asp Arg His Pro Asn Ile Gly Asp Gly Ala Leu Leu  
 225 230 235 240

gga gca tgt gtg act ata ctt ggt aac att aag ata ggc gct gga gca 768  
 Gly Ala Cys Val Thr Ile Leu Gly Asn Ile Lys Ile Gly Ala Gly Ala  
 245 250 255

atg gta gct gct ggt tgc ctt gtg tta aag gat gtt cct tgc cat agc 816  
 Met Val Ala Ala Gly Ser Leu Val Leu Lys Asp Val Pro Ser His Ser  
 260 265 270

atg gtg gct gga aat cca gca aaa ctc atc ggg ttt gtt gat gag caa 864  
 Met Val Ala Gly Asn Pro Ala Lys Leu Ile Gly Phe Val Asp Glu Gln  
 275 280 285

gat cca tct atg aca atg gag cat ggt gag tct tga 900  
 Asp Pro Ser Met Thr Met Glu His Gly Glu Ser  
 290 295 300

<210> 7  
 <211> 54  
 <212> ADN  
 <213> Séquence artificielle

<220>  
 <223> Description de la séquence artificielle:  
 oligonucléotide synthétique

<400> 7 54  
 gagagaggat cctctttcca atcataaacc atggcaacat gcatagacac atgc

<210> 8  
 <211> 46  
 <212> ADN  
 <213> Séquence artificielle

<220>  
 <223> Description de la séquence artificielle:  
 oligonucléotide synthétique

<400> 8 46  
 ggctcaccag actaatacac taaattgtgt ttacctcgag agagag

<210> 9  
 <211> 52  
 <212> ADN  
 <213> Séquence artificielle

<220>

<223> Description de la séquence artificielle:  
oligonucléotide synthétique

<400> 9  
gagagaggat cctcttatcg ccgcgttaat atgccaccgg ccggagaact cc 52

<210> 10  
<211> 45  
<212> ADN  
<213> Séquence artificielle

<220>  
<223> Description de la séquence artificielle:  
oligonucléotide synthétique

<400> 10  
gagccttacc agtctaattgt agtatatttc aacctcgaga gagag 45

<210> 11  
<211> 53  
<212> ADN  
<213> Séquence artificielle

<220>  
<223> Description de la séquence artificielle:  
oligonucléotide synthétique

<400> 11  
gagagaggat cccctcctcc tctcctcct atggctgcgt gcatcgacac ctg 53

<210> 12  
<211> 44  
<212> ADN  
<213> Séquence artificielle

<220>  
<223> Description de la séquence artificielle:  
oligonucléotide synthétique

<400> 12  
gctcaccagc ctaatacatt aaactttttc agctcgagag agag 44

<210> 13  
<211> 53  
<212> ADN  
<213> Séquence artificielle

<220>  
<223> Description de la séquence artificielle:  
oligonucléotide synthétique

<400> 13  
gagagaggat ccggccgaga aaaaaaaaaa atgttgccgg tcacaagtcg ccg 53

<210> 14  
<211> 49  
<212> ADN  
<213> Séquence artificielle

<220>  
<223> Description de la séquence artificielle:

oligonucléotide synthétique

<400> 14  
gagagaggat ccgacaagtt ggcataattt atgggtggatc tatcttcct

49

<210> 15  
<211> 43  
<212> ADN  
<213> Séquence artificielle

<220>  
<223> Description de la séquence artificielle:  
oligonucléotide synthétique

<400> 15  
cctgtgtgat tgctgtgtag tactctagaa actcgagaga gag

43

<210> 16  
<211> 67  
<212> ADN  
<213> Séquence artificielle

<220>  
<223> Description de la séquence artificielle:  
oligonucléotide synthétique

<400> 16  
gagagaggat ccgacaagtt ggcataattt atggcttgta taaacggcga gaatcgtgat 60  
ttttctt 67

<210> 17  
<211> 40  
<212> ADN  
<213> Séquence artificielle

<220>  
<223> Description de la séquence artificielle:  
oligonucléotide synthétique

<400> 17  
tacctcgtac cactcagaac tctagaaact cgagagagag

40

INSTITUT NATIONAL

de la

PROPRIÉTÉ INDUSTRIELLE

RAPPORT DE RECHERCHE  
PRELIMINAIREétabli sur la base des dernières revendications  
déposées avant le commencement de la rechercheFA 572814  
FR 9816163

DOCUMENTS CONSIDERES COMME PERTINENTS		Revendications concernées de la demande examinée
Catégorie	Citation du document avec indication, en cas de besoin, des parties pertinentes	
A	YOUSSEFIAN, S., ET AL. : "tobacco plants transformed with the O-acetyserine (thiol) lyase gene of wheat are resistant to toxic levels of hydrogen sulphide gas" THE PLANT JOURNAL, vol. 4, no. 5, 1993; pages 759-769, XP002078206 page 759 * abrégé; figure 1 *	1-40
A	SAITO ET AL: "modulation of cysteine biosynthesis in chloroplasts of transgenic tobacco overexpressing cysteine synthase (O-Acetylserine(thiol) -lyase)" PLANT PHYSIOLOGY, no. 106, 1 janvier 1994 (1994-01-01), page 887 895 XP002078205 ISSN: 0032-0889 * le document en entier *	1-40
A	WO 97 15673 A (CONSORTIUM ELEKTROCHEM IND ;LEINFELDER WOLFRED (DE); HEINRICH PETE) 1 mai 1997 (1997-05-01) pages 2,4,13; * figure 1; exemple 2; tableau 1B *	1-40
A	WO 98 55601 A (THORPE CATHERINE JANE ;KINNEY ANTHONY JOHN (US); RAFALSKI J ANTONI) 10 décembre 1998 (1998-12-10) * abrégé; figure 1 *	1-40
		DOMAINES TECHNIQUES RECHERCHES (Int.CL.6)
Date d'achèvement de la recherche		Examineur
17 septembre 1999		Holtorf, S
<p>CATEGORIE DES DOCUMENTS CITES</p> <p>X : particulièrement pertinent à lui seul Y : particulièrement pertinent en combinaison avec un autre document de la même catégorie A : pertinent à l'encontre d'au moins une revendication ou arrière-plan technologique général O : divulgation non-écrite P : document intercalaire</p> <p>T : théorie ou principe à la base de l'invention E : document de brevet bénéficiant d'une date antérieure à la date de dépôt et qui n'a été publié qu'à cette date de dépôt ou qu'à une date postérieure. D : cité dans la demande L : cité pour d'autres raisons &amp; : membre de la même famille, document correspondant</p>		

1

EPO FORM 1503 03.82 (P04C13)

DOCUMENTS CONSIDERES COMME PERTINENTS		Revendications concernées de la demande examinée
Catégorie	Citation du document avec indication, en cas de besoin, des parties pertinentes	
A	HOWARTH, J.R., ET AL. : "cysteine biosynthesis in higher plants; a new member of the Arabidopsis thaliana serine acetyltransferase small gene-family obtained by functional complementation of an Escherichia coli cysteine auxotroph" BIOCHIMICA ET BIOPHYSICA ACTA, vol. 1350, 1997, pages 123-127, XP002115632 * le document en entier *	1-40
A	ROBERTS, M.A., ET AL. : "cloning and characterization of an Arabidopsis thaliana cDNA clone encoding an organellar isoform of serine acetyltransferase" PLANT MOLECULAR BIOLOGY, vol. 30, 1996, pages 1041-1049, XP002115633 * le document en entier *	1-40
A	RUFFET, M-L., ET AL. : "subcellular distribution of serine acetyltransferase from Pisum sativum and characterization of an Arabidopsis thaliana putative cytosolic isoform" EUROPEAN JOURNAL OF BIOCHEMISTRY, vol. 227, 1995, pages 500-509, XP002115634 * le document en entier *	1-40
		DOMAINES TECHNIQUES RECHERCHES (Int.CL.6)
Date d'achèvement de la recherche		Examineur
17 septembre 1999		Holtorf, S
<p>CATEGORIE DES DOCUMENTS CITES</p> <p>X : particulièrement pertinent à lui seul Y : particulièrement pertinent en combinaison avec un autre document de la même catégorie A : pertinent à l'encontre d'au moins une revendication ou arrière-plan technologique général O : divulgation non-écrite P : document intercalaire</p> <p>T : théorie ou principe à la base de l'invention E : document de brevet bénéficiant d'une date antérieure à la date de dépôt et qui n'a été publié qu'à cette date de dépôt ou qu'à une date postérieure. D : cité dans la demande L : cité pour d'autres raisons &amp; : membre de la même famille, document correspondant</p>		

1

EPO FORM 1503 03.82 (P04C13)

**ANNEXE AU RAPPORT DE RECHERCHE PRELIMINAIRE  
RELATIF A LA DEMANDE DE BREVET FRANÇAIS NO.**

FA 572814  
FR 9816163

La présente annexe indique les membres de la famille de brevets relatifs aux documents brevets cités dans le rapport de recherche préliminaire visé ci-dessus.  
Lesdits membres sont contenus au fichier informatique de l'Office européen des brevets à la date du  
Les renseignements fournis sont donnés à titre indicatif et n'engagent pas la responsabilité de l'Office européen des brevets,  
ni de l'Administration française

17-09-1999

Document brevet cité au rapport de recherche	Date de publication	Membre(s) de la famille de brevet(s)	Date de publication
WO 9715673 A	01-05-1997	DE 19539952 A	30-04-1997
		CA 2235752 A	01-05-1997
		CN 1200764 A	02-12-1998
		CZ 9801269 A	15-07-1998
		EP 0858510 A	19-08-1998
		HU 9900078 A	28-04-1999
		PL 327187 A	23-11-1998
WO 9855601 A	10-12-1998	AU 7727098 A	21-12-1998